Cognitive Imprecision and Stake-Dependent Risk Attitudes*

Mel Win Khaw
Microsoft Corporation

Ziang Li
Imperial College London

Michael Woodford
Columbia University

May 10, 2026

Abstract

In an experiment that elicits subjects' willingness to pay (WTP) for the outcome of a lottery, we document a systematic effect of stake sizes on the magnitude and sign of the relative risk premium, and find that there is a log-linear relationship between the monetary payoff of the lottery and WTP, conditional on the probability of the payoff and its sign. We account quantitatively for this relationship, and the way in which it varies with both the probability and sign of the lottery payoff, in a model in which all departures from risk-neutral bidding are attributed to an optimal adaptation of bidding behavior to the presence of cognitive noise. Moreover, the cognitive noise required by our hypothesis is consistent with patterns of bias and variability in judgments about numerical magnitudes and probabilities that have been observed in other contexts. We thus provide foundations for the kind of nonlinear distortions in lottery valuation posited by prospect theory, that we believe can provide an interpretation for the observed instability across contexts of estimated prospect-theoretic parameters.

*Listed affiliations are current ones; Dr. Khaw's work on this project took place while he was a postdoctoral research affiliate of Columbia University. We thank Rahul Bhui, Helga Fehr-Duda, Cary Frydman, Paul Glimcher, Lawrence Jin, Eric Johnson, Ryan Oprea, Dragan Rangelov, Charlie Sprenger, Peter Wakker and George Wu for helpful comments, and the National Science Foundation for research support.

One of the more puzzling features of decision making under risk in the laboratory is the fact that the same experimental subjects can display either risk-averse or risk-seeking behavior, depending on the nature of the choices presented to them (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Kahneman and Tversky stress in particular the existence of a "fourfold pattern of risk attitudes," according to which the sign of the risk premium reverses when one considers random losses rather than random gains, or when one considers a gamble with only a low probability of a non-zero outcome rather than one with a relatively high probability of the most extreme outcome.

But there is another way in which apparent risk attitudes vary with the terms of a gamble. A literature beginning with Markowitz (1952) has argued that scaling up the stakes involved in a gamble (while keeping the probability of a non-zero payoff the same) moves subjects toward greater risk-aversion (or less risk-seeking) in the case of valuation of a random gain, while a similar increase in the stake size moves them toward greater risk-seeking (or less risk-aversion) in the case of valuation of a random loss.[1] These effects are illustrated in Figure 1.

Here we plot data from an experiment that elicits the amount that subjects are willing to pay (WTP) for the outcome of a lottery $(X; p)$, that pays monetary amount X with probability p, and zero otherwise.[2] Figure 1 plots the average value (across trials, for the median subject) of WTP as a multiple of the payoff X, and plotted as a function of p, the probability of the non-zero payoff;[3] the sign of the risk premium is indicated by whether WTP/X is above or below the dotted diagonal line. For both of the two stake sizes shown, we verify the fourfold pattern of Tversky and Kahneman. But the size of $|X|$ also shifts the relationship, both in the case of lotteries involving random gains (the left panel) and ones involving random losses (the right panel). For each value of p and both signs of X, the shifts have the signs summarized above.

Below we further document not only the sign of these stake-size effects but that they are roughly *log-linear:* letting EV denote a lottery's expected value, we show that the relative risk premium $\log(WTP/EV)$ is a decreasing affine function of $\log|X|$, in both the gain and loss domains, and for each value of p. Moreover, the size of the negative elasticity depends on p: it is larger the smaller the probability of the non-zero payoff (whether it is a gain or a loss). These are all features of measured stake-size effects that we wish to understand.

A recent literature has argued that the fourfold pattern of Kahneman and Tversky can be understood as an efficient response to imprecision in people's understanding of how a change in the probability of a lottery's paying off should affect the value of the lottery.[4] Essentially, the idea is that if decisions must be based on a noisy mental representation of the frequency with which the lottery pays off, rather than the exact probability stated by the experimenter, an optimal bidding rule will also take into account the prior probability with which lotteries generally pay off in that environment, leading to valuations that would

[1] See also Hershey and Schoemaker (1980), Kachelmeier and Shehata (1992), Fehr-Duda *et al.* (2010), Scholten and Read (2014), and Bouchouicha and Vieider (2017).

[2] The experiment is described in more detail in section 1 below.

[3] The data plotted here are for the 15 subjects ("group 5") who all face the same set of 640 lotteries. The stake-size effects for these subjects are compared with those of our other subjects in Appendix section F.1.

[4] See in particular Khaw *et al.* (2021), Enke and Graeber (2023), Oprea (2024), Vieider (2024), Bedi *et al.* (2025), and Bouchouicha *et al.* (2025b).

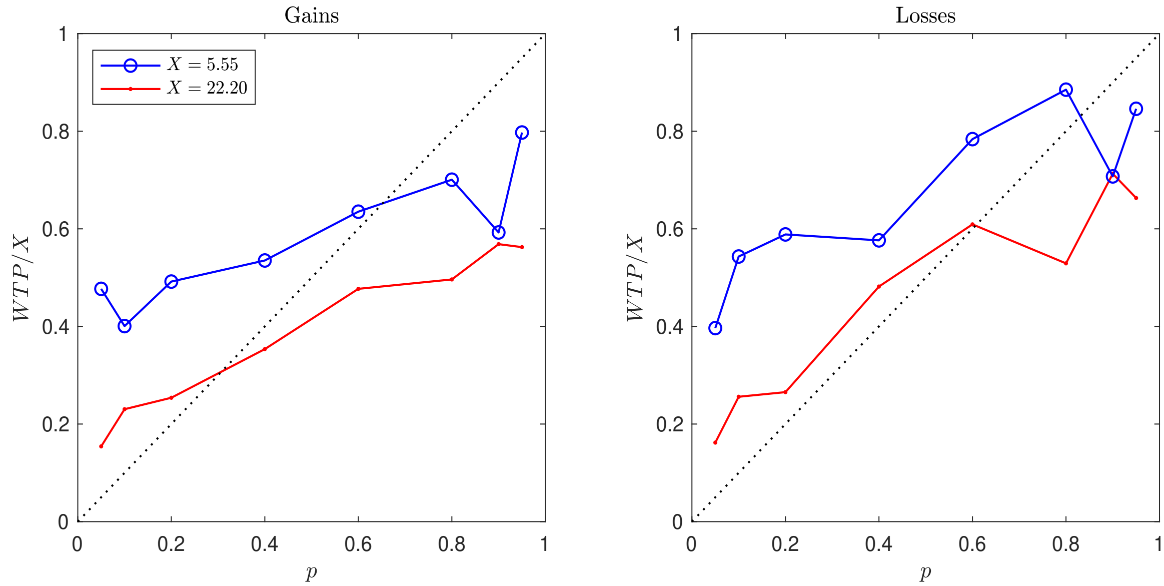


Figure 1: Illustration of the consequences for measured risk premia of increasing the payoff size $|X|$ in all lotteries by a factor of 4. (See text for discussion.)

be appropriate for a larger value of p than the actual one when the actual value of p is small, and valuations that would be appropriate to a smaller value of p than the actual one when the actual value is large. In this view, the fourfold pattern is simply another example of a more general phenomenon of "behavioral attenuation" in response to cognitive uncertainty that is observed in a wide range of contexts (Enke *et al.*, 2026).

Here we ask whether stake-size effects on lottery valuations of the kind just described can also be explained as an efficient response to imprecision in subjects' mental representation of the characteristics of the lotteries that they must evaluate. This requires an extension of two kinds to the explanations that have been offered for the fourfold pattern. First, we wish to consider the separate effects on measured relative risk premia of both changes in the probability p and changes in the size $|X|$ of the monetary payoff. Because the experiments of Enke and Graeber (2023) and Oprea (2024) involve only variations in p and the sign of the payoff, keeping $|X|$ the same across trials, it remains ambiguous whether the fourfold pattern might be attributed to bids based on a noisy representation of the lottery's value (computed on the basis of a precise understanding of its terms), rather than having to compute an value estimate from a noisy representation of the probability itself. We instead conduct an experiment in which both $|X|$ and p are independently varied, and consider whether the patterns of departure from risk-neutral valuation are better fit by a model in which there is independent noise in separate representations of $|X|$ and p than by a model in which there is only noisy retrieval of an exactly computed lottery value.

And second, we are interested in explaining specific features of the stake-size effects that we measure: that they are log-linear, and that the negative elasticity is greater the smaller is p. Predictions regarding these questions depend on specific quantitative hypotheses about the imprecise representations of monetary payoffs that subjects draw upon in making their bids. We motivate these hypotheses by reference to evidence on the nature of imprecise mental number representations from another domain: approximate judgments about the number of items in a visual display.

We review evidence about the nature of the "approximate number system" from studies of this kind, and show evidence for not only an imprecise sense of the sizes of numbers, but for characteristic biases in number estimates based on ANS representations. These biases are strikingly similar to the biases in the way that stake size affects lottery valuations in our experiment. Biases in estimates of the numerosity of visual arrays are found to be approximately log-linear, and of the sign needed to account for the stake-size effects that we document. Moreover, the biases are context-sensitive in ways that indicate that the precision of number representation responds to changes in the incentive for precision; numerosity estimation experiments provide support for a particular cost function for precision that we incorporate into the model of lottery valuations that we propose.[5]

This model predicts that the size of the lottery payoff $|X|$ should be represented with less precision on trials where p is smaller, because there is a smaller financial incentive for accurate recognition of the exact size of the payoff on these trials. It is then an optimal adaptation to this reduced precision for the subject's bid to be less sensitive (in percentage terms) to variation in the internal representation on these trials. The result is that $\log(WTP/EV)$ should be a more sharply decreasing function of $\log |X|$ on the low-p trials, as we observe.

The paper proceeds as follows. In section 1, we present our experiment and document the existence of log-linear stake-size effects, the strength of which varies with p. Section 2 then explains the key mechanisms of our model of variable stake-size effects, reviewing the ways in which our model directly builds upon quantitative models of imprecise number representation. Section 3 discusses additional details of the quantitative model that we fit to our experimental data, and section 4 then discusses the fit of the model with the data moments reported in section 1. Section 4 also compares the fit of our baseline model with that of a variety of alternative specifications of the cognitive noise, to clarify the quantitative importance of each of several different types of noise present in the complete model. Section 5 offers a concluding discussion, comparing the cognitive-noise interpretation of measured risk attitudes with the way these are modeled in prospect theory.

1 Determinants of Risk Attitudes in a Lottery-Valuation Experiment

We begin by documenting the stake-dependence of risk attitudes through a new experimental study. As in previous studies following Tversky and Kahneman (1992), we elicit values for lotteries that are described to experimental subjects, and map out the complete fourfold pattern of risk attitudes by presenting lotteries involving both gains and losses, and both large and small values of p.[6] But in this experiment, we also consider a range of different

[5]Our model resembles theories of "rational inattention" (Sims, 2003), in that we assume that precision is optimized for a specific context, positing a cost of precision that can be traded off against other objectives (here, the expected financial consequences of inaccurate bidding). But unlike much work in that literature, we base our assumed cost function on empirical studies of perceptual accuracy.

[6]Our experimental procedure differs from that of Tversky and Kahneman (1992) in that they elicit certainty-equivalent (CE) values using multiple-price lists, while we elicit subjects' willingness-to-pay using a BDM auction, as discussed further below. But as discussed in Appendix section H, the log-linear stake-size

stake sizes $|X|$ for each value of p, and vary the stake size from trial to trial in a way that is statistically independent of the choices for p and the sign of X. We use a larger number of stake sizes than most previous studies that vary stake size at all, allowing us to show not only that stake-size effects exist, but that they are roughly log-linear.

Our study differs from most previous work in another important respect as well. Many studies elicit only a single valuation from each subject for a given lottery; it is assumed that a given lottery should have a well-defined value under a given person's preferences, and that one need only ask once to elicit it. Our theoretical framework assumes instead that responses to a given decision problem should vary randomly from trial to trial, owing to noise in the internal representation of the problem; and we are furthermore interested in measuring this randomness, because our model of optimal adaptation to cognitive noise implies that biases in the average valuation of a given lottery depend on the degree of noise in the internal representation. To test such a theory, we need to fit the predictions of our model to data on *both* average valuations and the degree of trial-to-trial variation in these valuations. Hence (as in Khaw *et al.,* 2021) we adopt an experimental procedure from studies of the accuracy of perceptual judgments, in which a variety of items are presented to the subject for evaluation in random order, with the same item (in our case, the same lottery) appearing multiple times over the course of the experiment.[7]

1.1 Experimental Design

A total of 28 subjects[8] participated in an experiment in which they were required to bid dollar amounts that they were willing to pay to obtain the outcome of a lottery which would pay an amount X with a probability p, and otherwise zero. The screen interface is shown in Figure 2. On each trial, the lottery offered is visually represented by a two-color vertical bar, the two segments of which represent the two possible outcomes. The probability of each outcome is indicated by a two-digit number inside that segment of the bar (showing the probability of that outcome in percent); the relative probabilities of the two outcomes are also indicated visually by the relative lengths of the two differently-colored segments. The monetary payoffs associated with each outcome (X and 0 respectively) are indicated by numbers at the two ends of the bar. (Note that the probabilities of *both* outcomes are displayed to the subject, with each given equal prominence, though to simplify notation we refer to the probabilities in any given decision problem by specifying only the probability of the non-zero payoff.)

A wide range of values of the probability p were used on different trials, corresponding to the different columns in Figures 3 and 4.[9] Five different values of the non-zero payoff

effects that we find are also present in the data of Gonzalez and Wu (2022), who elicit CE values.

[7] An early example of the use of this method in the study of risk preferences is the work of Mosteller and Nogee (1951).

[8] These were student subjects recruited at Columbia University, following procedures approved by the Columbia Institutional Review Board under protocol IRB-AAAQ2255. For additional details of our procedures, see Appendix section A and the online document "Supplementary Information: Experimental Procedure."

[9] Each subject faced up to 8 of these values of p, but not the complete set, so as to allow more repetitions of the lotteries presented to a given subject. The particular values of p used with different groups of subjects are explained in the Appendix, section A.3.

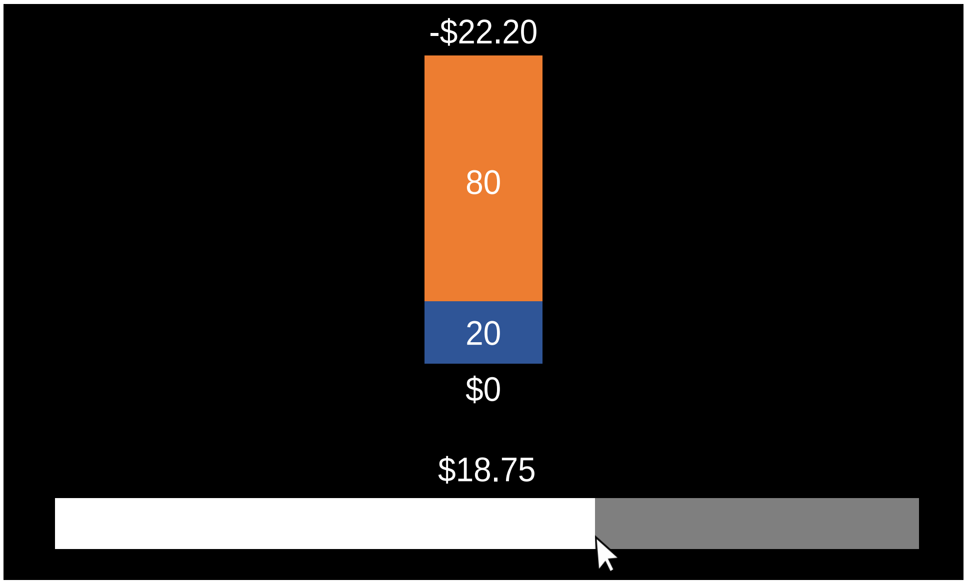


Figure 2: Example of the screen seen by an experimental subject.

were used: \$5.55, \$7.85, \$11.10, \$15.70, and \$22.20. (These values were chosen to be roughly equal distances apart along a logarithmic scale; we did not use integer numbers of dollars, so as not to encourage subjects to treat the problem as a test of arithmetic ability.) Each of these payoffs could be either positive (a possible gain) or negative (a possible loss); thus on a given trial, X could be either \$22.20 or -\$22.20 (as in the case shown in Figure 2). Each of the possible values of p was paired with all ten of the possible values of X (both positive and negative), and the same decision problem $(X; p)$ was presented to any given subject 8 times over the course of the experimental session, but with the problems randomly interleaved.

On each trial, after presentation of the lottery, the subject was required to indicate the amount that they were willing to pay for the outcome of the lottery, by moving a slider in a horizontal bar using the computer mouse. In the case of a lottery involving losses, the subject had to indicate the amount that they were willing to pay to *avoid* having to pay the outcome of the lottery. Thus in our discussion below, we refer to the subject's bid as WTP, their declared willingness-to-pay.[10] As shown in Figure 2, the dollar bid implied by a given slider position was shown on the screen. We used this method of elicitation of subjects' valuations, rather than the commonly used multiple-price-list procedure, because it allowed subjects to give a precise response rather than only indicating an interval. The fact that subjects' responses were not exactly the same on multiple repetitions of the same decision problem is not a disadvantage of the procedure in our case; the variability of trial-by-trial responses is actually one of the things that we wish to measure, rather than being regarded as a nuisance. Subjects' choices were incentivized by selecting one of their trials at random at the end of the experiment to be the one that mattered, and then conducting a BDM auction (Becker, DeGroot, and Marschak, 1964) in which the subject's bid on that trial was compared with a random bid chosen by the computer (independent of the subject's bid).[11]

[10]In the case of a lottery involving losses, we define WTP as the negative of the amount indicated by the subject's slider, so that in all cases WTP represents an elicited certainty-equivalent value of the lottery.

[11]The incentives created by this procedure are discussed further in the Appendix, sections A.1 and A.2.

On some trials, subjects submit a bid of zero (the leftmost position of the slider).[12] Since a subject should never be genuinely indifferent between the lottery offered and zero for sure (the lottery either clearly dominates zero, in the case of a random gain, or is clearly dominated by zero, in the case of a random loss), we interpret these responses as a subject declining to bid, rather than a considered bid that happens to be equal to zero. The trials on which the subject declines to bid are discarded in the analysis below of subjects' willingness-to-pay. (See section 3.3 for our theoretical interpretation of the zero-bid trials.)

1.2 Results

Figures 3 and 4 present statistics regarding subjects' reported willingness-to-pay (WTP) for each of 110 different lotteries: 11 different values of p (the eleven columns), and 5 different values of $|X|$ (the horizontal axis of each panel), in both the case of random gains (the top panel of each column) and the case of random losses (the lower panel of each column). For each lottery, subjects' bids are described in terms of the implied value of $\log(WTP/EV)$, where the expected value of the lottery is given by $EV = pX$. This can be interpreted as a measure of the relative risk premium in the case of lotteries involving random losses; the negative of this quantity measures the relative risk premium in the case of lotteries involving gains. Risk-neutral valuations (or perfectly accurate bidding, given the objective assumed in equation (A.1) below) would correspond to a value of zero on every trial, for each lottery $(X; p)$. Thus the statistics presented in the figures measure the degree of discrepancy with respect to this benchmark, for those trials on which the subject submits a (non-zero) bid.

In the case of each lottery, the dot indicates the median value (across the 28 subjects) of the subject-level mean $\log(WTP/EV)$. The vertical whiskers mark an interval $\pm s$ around the mean, where s is the median value of the subject-level standard deviation of $\log WTP$.[13] Note that we follow authors like Tversky and Kahneman (1992) in stressing (and fitting our models to) the subjects' median behavior, without modeling the outliers.[14] But unlike many studies in that tradition, however, we are interested in the degree of subject-level response variability, across repeated presentations of the same lottery to the same subject, which we use to identify the degree of noise in the "average subject's" cognitive processes.

The solid line in each panel indicates the prediction of an OLS regression model (with separate coefficients for each panel), the "general affine model" discussed further below. Figure 3 shows the distributions of bids in the case of lotteries with relatively low values of p (between 0.05 and 0.40), while Figure 4 shows the corresponding distributions in the case of larger values of p (between 0.50 and 0.95). The dotted horizontal line in each panel indicates the benchmark of risk-neutral valuation ($WTP = EV$).

The importance of presenting choices involving real as opposed to merely hypothetical payoffs, especially for the measurement of stake-size effects, is illustrated by Holt and Laury (2002, 2005).

[12]This occurs about 1.2 percent of the time overall, though more frequently when the EV of the lottery is small. See the Appendix, section A.4, for more information about these bids.

[13]See Appendix section E.2 for further details of the computation of the data moments.

[14]Thus our hypothesis of optimal adaptation to the variance of cognitive noise should be understood to be a hypothesis that the *median* responses are close to being optimal, rather than that every individual subject's responses must be. It is a "wisdom of the crowd" hypothesis, like Muth's (1961) "rational expectations hypothesis."

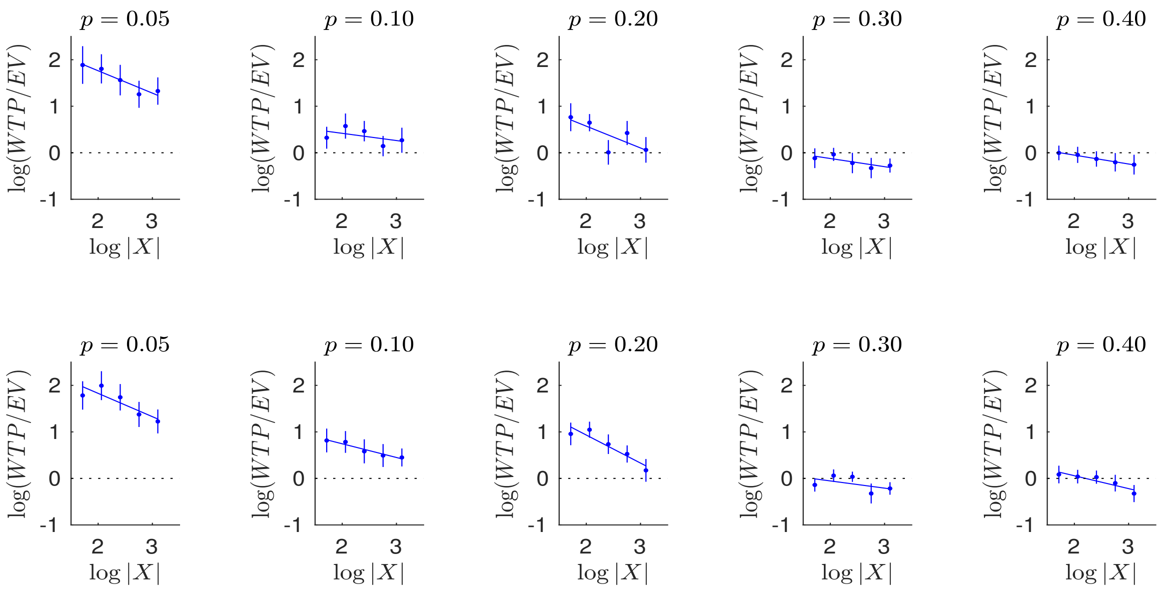


Figure 3: The distribution of values for WTP as a multiple of EV, for lotteries with different values of p (the different columns) and $|X|$ (the horizontal axis within each panel). The top panel in each column refers to lotteries involving random gains ($X > 0$), and the bottom panel to lotteries involving random losses ($X < 0$).

Several features of our data are immediately evident from these figures. First, we see that our experiment confirms the fourfold pattern of risk attitudes documented by Tversky and Kahneman (1992): subjects' bids are for the most part risk-averse in the case of risky gains when p is 0.30 or larger ($0 < WTP < EV$), and in the case of risky losses when p is 0.10 or less ($WTP < EV < 0$), but are instead mostly risk-seeking in the case of risky gains when p is 0.10 or less ($0 < EV < WTP$), and in the case of risky losses when p is 0.30 or larger ($EV < WTP < 0$).

Yet in addition, we also see a consistent stake-size effect: in each of the 22 panels, the geometric mean value of WTP/EV becomes smaller (or at least becomes no larger) the larger the value of $|X|$. In the transitional case (with respect to the Tversky-Kahneman pattern) where $p = 0.2$, this means that for small stake sizes we observe risk-seeking bidding in the gain domain but risk-averse bidding in the loss domain, while for larger stake sizes we instead observe risk-averse bidding in the gain domain and risk-seeking bidding in the loss domain (the "alternative fourfold pattern" of Scholten and Read, 2014).[15] But the sign of the stake-size effect is the same (in both the gain and the loss domains) for all of the other values of p as well, though stake-size effects are most dramatic in the case of the smallest values of p (as is consistent with previous findings).

We also observe that the stake-size effects in each panel are approximately log-linear: the mean value of $\log(WTP/EV)$ for each lottery comes close to falling on the regression line for

[15] The same alternative fourfold pattern is observed, though in a less pronounced way, when $p = 0.4$, since in this case the mean relative risk premium changes sign for the smallest value of $|X|$. We can also observe the relative risk premium changing sign, in the direction predicted by the alternative fourfold pattern, for the cases $p = 0.1$ and $p = 0.25$ in the data of Gonzalez and Wu (2022). See Figure 12 in Appendix section H.

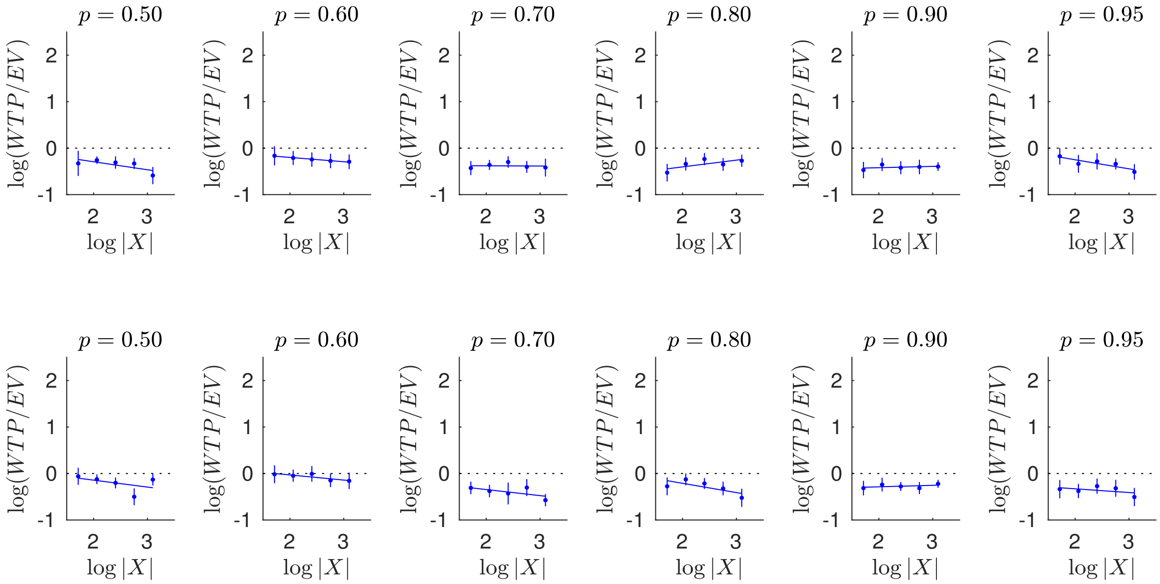


Figure 4: The same information as in Figure 3 (and using the same format), but now for probabilities $p \geq 0.50$.

that panel, meaning that (fixing p and the sign of X) mean $\log(WTP/EV)$ is a decreasing linear function of $\log|X|$. Moreover, not only is the slope of this linear relationship negative (or at least non-positive), it is never more negative than -1, so that increasing the stake size (for given p) increases the mean $\log|WTP|$, as one might expect.

Finally, we note that subjects' bids for the same lottery vary from trial to trial; this within-subject variability of responses is especially notable when the probability p of the non-zero payoff is small. In Figure 5, we show how the standard deviation of $\log WTP$ across trials varies as both p and $|X|$ increase; each of the horizontal segments reports the variability of bids for the median subject for the five different payoff magnitudes $|X|$, for a given choice of p and of the sign of the payoffs.[16] The values of $\log WTP$ are plotted with the mean value of $\log WTP$ for that lottery as the horizontal coordinate; this allows us to verify that the larger standard deviation of $\log WTP$ does not simply reflect the possibility that a given amount of variance in the dollar bids reported (e.g., due to motor noise in adjusting the slider) will appear as greater variation in percentage terms in the case of smaller bids. We see that even when the average size of bids on two lotteries is similar, the bids are more variable across trials in the case of lotteries with low values of p. This is worth noting because, as is evident in Figure 3, stake-size effects are also largest when p is small; and under the theory that we propose, it is not an accident that the two phenomena are strongest in the same cases.

[16]For the sake of legibility, we include in each panel the segments for only five values of p; the values chosen for display are such that each successive increase in p increases $\log p$ by the same amount. We also use in this figure only the data from the same 15 subjects as in Figure 1 ("group 5" in Appendix section A.3), so that the comparisons across different values of p represent data from the same set of subjects in each case.

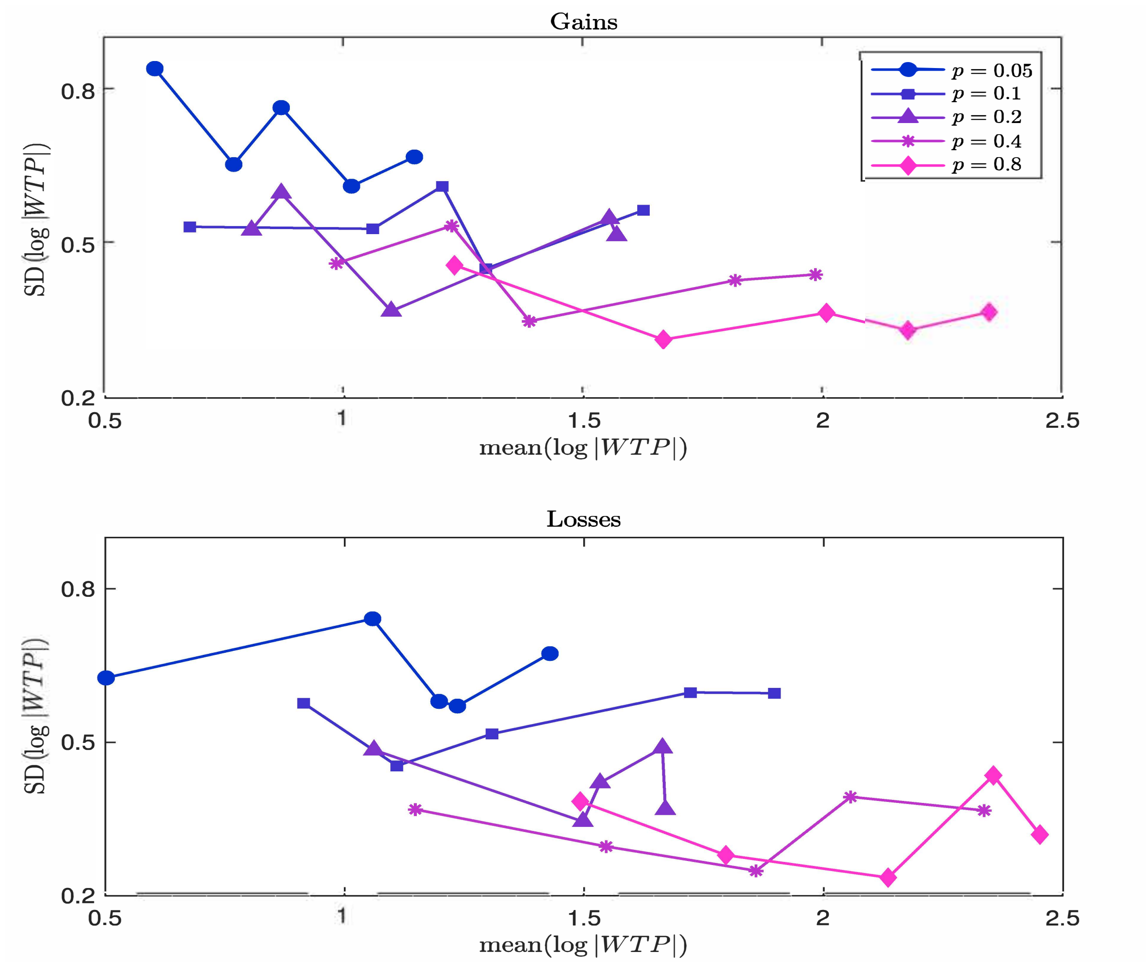


Figure 5: Trial-to-trial variability of bids for a given lottery, in the case of the median 640-trial subject.

1.3 Log-Linear Stake-Size Effects

Visual inspection of Figures 3 and 4 suggests a downward-sloping log-linear relationship between WTP/EV and the size of the monetary payoff X in each of the panels, and moreover that this relationship is essentially the same regardless of the sign of X. Here we present statistical evidence that this is indeed an accurate characterization of our average subject's responses. We distinguish between a series of progressively more restrictive statistical models of our subjects' behavior. In the most general (purely atheoretical) characterization of the data, we suppose that for each lottery $(X; p)$ there is a distribution of values for the willingness-to-pay of the form

$$\log \frac{WTP}{EV} \sim N(m(p,X),\, v(p,X)). \tag{1.1}$$

In what we call our "unrestricted model," there are thus two parameters, $m(p,X)$ and $v(p,X)$, to be estimated for each lottery, with no restrictions linking the parameters for any given lottery to those for any other lotteries. Our "symmetric model" instead imposes the restrictions $m(p, X) = m(p, -X)$ and $v(p, X) = v(p, -X)$, so that the distribution of values for WTP/EV depends only on p and $|X|$: it is the same for random losses as for random gains.

Alternatively, we can restrict the general model by assuming that for any p and any sign of X, $m(p, X)$ be an affine function of $\log |X|$. Our "general affine model" allows the slope and intercept for each value of p differ depending whether gains or losses are involved; this is the characterization of the data assumed in fitting the regression lines shown in each of the panels of Figures 3 and 4. Our "symmetric affine model" imposes all of the restrictions of both the symmetric model and the general affine model, so that

$$m(p, X) = \alpha_p + \beta_p \log |X|, \tag{1.2}$$

regardless of the sign of $|X|$, for coefficients (α_p, β_p) that depend only on the value of p. The "bounded symmetric affine model" imposes all of these restrictions, plus the further restriction that $-1 \leq \beta_p \leq 0$ for all p.

We also consider a family of models that impose even tighter restrictions on the values of the $\{\beta_p\}$. For each possible threshold p^*, we consider a model that imposes all of the restrictions of the bounded symmetric affine model, and in addition requires that $\beta_p = 0$ for all $p \geq p^*$. The most restrictive case is the "no stake effects" model that requires that $\beta_p = 0$ for all p. Consideration of more restrictive models in which β_p is required to equal zero for all large enough p allows us to obtain quantitative measures of the importance of allowing for stake effects in order to match our data.

Finally, we consider models in which $m(p, X)$ is allowed to be a non-linear function of $\log |X|$, but still one with fewer free parameters than our (otherwise unrestricted) "symmetric model." Specifically, we consider special cases of the symmetric model in which $m(p, X)$ is a quadratic or cubic function of $\log |X|$; these are the models called "symmetric quadratic" and "symmetric cubic" in Table 1.[17]

Table 1 reports measures of the goodness of fit of each of these models to the data on the distribution of bids of the average subject. Given that each of the models assumes a log-normal distribution of responses (1.1), the likelihood of the data under any specification of the model parameters is a function of 220 data moments: the quantities $(\hat{m}_j, \hat{v}_j)$ for each of the 110 possible lotteries (p_j, X_j). Here for each lottery j, $\hat{m}_j$ is the mean and $\hat{v}_j$ the variance of the sample distribution of values for $\log(WTP/EV)$. The values of these moments that we attribute to the average subject are the ones plotted in Figures 3 and 4. The likelihood also depends on N_j, the number of trials on which lottery j is evaluated. (See the Appendix, section E, for further details.) The parameters of each model are chosen to maximize the likelihood of these data moments, subject to the restrictions specified above.

The first column of the table reports the maximized value of the log likelihood (LL) for each model. As one would expect, each successive additional restriction on the model reduces the optimized value of LL. The second column instead reports the value of the Bayes Information Criterion (BIC) for each model, defined as $BIC \equiv -2LL + \sum_k \log N_k$, where for each free parameter k of the model, N_k is the number of observations for which parameter k is relevant.[18] This is a measure of goodness of fit which (unlike LL alone) penalizes the use

[17]To reduce the size of the table, we do not also present statistics for the asymmetric variants of these models, but only the ones that assume a common relationship in both gain and loss domains.

[18]See, for example, Burnham and Anderson (2002), p. 271.

Model	# params	LL	BIC	K
unrestricted model	220	-1503.0	3334.7	1
symmetric model	110	-1525.5	3297.9	9.8×10^7
symm. cubic	99	-1526.7	3347.6	0.0016
symm. quadratic	88	-1529.6	3312.3	73,000
general affine model	154	-1513.2	3331.0	6.4
symmetric affine model	77	-1531.7	3275.3	7.9×10^{12}
bounded symm. affine	76	-1531.8	3271.2	6.1×10^{13}
$\beta_p = 0$ for $p \geq 0.5$	71	-1534.0	3257.7	5.3×10^{16}
$\beta_p = 0$ for $p \geq 0.3$	69	-1537.0	3256.7	8.7×10^{16}
no stake effects	66	-1546.8	3264.5	1.8×10^{15}
cognitive noise model	3	-1602.5	3236.3	2.3×10^{21}

Table 1: Measures of the goodness of fit of alternative statistical models of the average subject's responses. For each model, the number of free parameters (# param), the log likelihood (LL), and the Bayes Information Criterion (BIC) are reported, as well as the Bayes factor K by which each model is preferred to the unrestricted model.

of additional free parameters, making it possible for a more restrictive model to be judged better (as indicated by a lower BIC). The final column provides an interpretation of the BIC differences between the different models, by reporting the implied Bayes factor K by which the model in question should be preferred to the unrestricted model (used as the baseline).[19]

While the log likelihood is lower for more restrictive versions of the model, the BIC can also be lower, if the greater parsimony of the more restrictive model outweighs the somewhat poorer fit to the individual data moments. This is what we find when we move from the unrestricted model to the bounded symmetric affine model: while LL is reduced (by 28.8 log points), the BIC nonetheless falls (by 63.5), corresponding to a Bayes factor in favor of the more parsimonious model of more than 60 trillion. This is also a lower BIC (and correspondingly a larger Bayes factor) than in the case of any of the less-restricted models, such as the general symmetric model, or the quadratic or cubic models.[20] Thus our data are more consistent with a characterization of the form assumed by the bounded symmetric affine model.

When we consider additional restrictions on the β_p coefficients, we find that the BIC can be further reduced (and the Bayes factor corresponding increased) by imposing the restriction $\beta_p = 0$ for all large enough values of p; this is illustrated in the table for the cases in which the cutoff probability is either 0.3 or 0.5, and also for the case in which we require β_p to be zero for all p. The largest Bayes factor is obtained if we set $\beta_p = 0$ for all $p \geq 0.3$. The fact that the Bayes factor is larger for this model than for the one with no stake-size effects (as indeed would also be true in the case of a higher cutoff, such as 0.5) means that we do find statistically significant stake-size effects, with the same sign as those reported by authors in

[19]The Bayes factor K in favor of model M_2 over model M_1 is given by $\log K = (1/2)[\text{BIC}(M_1) - \text{BIC}(M_2)]$. See Burnham and Anderson (2002), p. 303.

[20]The cubic does not even have as low a BIC as the unrestricted model: though it is more parsimonious, the reduction in the log-likelihood owing to the restrictions outweighs the reduction in the penalty for free parameters. As a result, the Bayes factor in favor of the restricted model is less than 1 in this case.

the tradition started by Markowitz (1952) — i.e., that WTP/EV is a decreasing function of $|X|$, at least in the case of all small enough values of p.

Thus the best atheoretical characterization of our data, among those considered here, is one in which WTP/EV is a log-linear decreasing function of $|X|$, with a slope that depends on p and is most clearly negative in the case of low values of p. This relationship is essentially the same regardless of whether the lotteries involve gains or losses, and the elasticity β_p is always between 0 and -1. We show below that all of these regularities are predictions of a model of optimal bidding in the presence of cognitive noise.[21]

2 Imprecise Number Representation and Stake-Size Effects

We now explain how the departures from risk-neutral valuations documented in the previous section can be explained as a consequence of subjects' responses being based on an imprecise mental representation of the properties of the lottery that they face on a given trial, rather than upon its actual (exact) characteristics. In particular, the existence of log-linear stake-size effects that are strongest when the probability of a lottery's paying off is low, can be understood as an optimal response to the nature of the imprecision in mental representation of the size of the monetary payoff offered by a given lottery. The model of imprecision that we rely upon is furthermore based on known properties of noisy semantic representations in the brain of numbers more generally. We review here some properties of imprecise number representation that are observed in other contexts, as motivation for the model that we propose.

2.1 Consequences of a Logarithmic "Mental Number Line"

It is observed that people have a sense of the size of numbers that can be used to make approximate quantitative judgments (for example, about whether adding five objects to nine objects should give one a larger or smaller number of items than are contained in a group of twelve objects), without refering to memorized facts of arithmetic or other exact operations (such as counting). This "approximate number system" (ANS) seems to develop in babies earlier than language (and hence the use of number words, let alone training in arithmetic), and is found to be present in remote tribes whose language lacks many number words, and who have not attended school.[22]

Even adults who have been formally schooled continue to have access to imprecise "semantic" representations of the size of numbers, including numbers presented to them using number words or symbols. The brain region (in the intraparietal sulcus, IPS) in which neural representations of number information are found shows similar patterns of activation a number is displayed visually (presentation of an array containing that number of objects) and when the same number is indicated using a number word or symbol (Piazza *et al.*, 2007). Hence even when number information is presented using symbols, an imprecise semantic

[21] The log likelihood and corresponding BIC statistic for that structural model are also shown in Table 1, on the bottom line. See further discussion below.

[22] See Anobile *et al.* (2016), Dehaene (2011), Nieder (2016), Odic and Starr (2018), and Piazza (2010) for reviews.

representation is created in the brain, along with the recognition of the exact number in the parts of the brain involved in language use. This imprecise ANS representation can be accessed more quickly and easily, and is particularly likely to be drawn upon when judgments about numbers must be made rapidly (Moyer and Landauer, 1967; Dehane *et al.*, 1990), or when numbers must be recalled from memory (Dehaene and Marques, 2002).

Our hypothesis here is that subjects in our experiment make intuitive judgments about the value of a lottery $(X; p)$ on the basis of imprecise semantic representations of the quantities X and p. Our emphasis in this section is on the nature of the imprecise representation of the size $|X|$ of the monetary payoff, which we hypothesize to be noisy in a way that is similar to the ANS representation of numbers in other contexts.

Evidence suggesting that ANS number representations may be drawn upon in choices under risk is provided by Barretto-Garcia *et al.* (2023) and De Hollander *et al.* (2025a), who show that differences in the precision with which different experimental subjects represent the numerosity of visual arrays — measured both behaviorally (in a pure numerosity-discrimination task) and through brain scans— correlate with both differences in the randomness of the same subjects' choices under risk and their degree of departure from risk-neutrality in those choices. De Hollander *et al.* (2025b) even provide evidence of a causal connection, showing that larger departures from risk-neutral valuations can be produced by perturbation of brain activity in regions of the parietal cortex known to be involved in number representation.

We should be clear that an explanation of this kind doesn't assume that the DM *doesn't precisely understand* the monetary amount that they have been told — it is after all still visible on the screen while they select their bid. We assume however that they select their bid through an intuitive process that uses the noisy representation r_x rather than the exact number X as an input. And we suppose that the process through which the noisy representation r_x is generated and then used to select a bid that seems appropriate is not accessible to the DM's conscious awareness. While they know the number X that is on the screen, they aren't able to compare the noisy representation r_x to the exact number X (and recognize that on a particular occasion the noisy representation will lead to an over- or under-estimate); nor do they see the calculation through which the representation r_x leads to a particular bid, so that it should be possible to correct the calculation using X rather than r_x as the input.[23]

The nature of ANS representation of numbers has been studied most extensively in the case of people's ability to perceive the approximate number of items in a visual scene, say an irregular array of dots. A number of quantitative regularities are well-documented.[24] When subjects are asked whether one array contains a larger or smaller number of items than another, the probability of a correct response is increasing in the difference in the two numbers. Moreover, the difference required to produce a given degree of accuracy is larger the larger the two numbers are; in fact, it is often found that it is essentially the *percentage* difference in the two numerosities that determines the probability that one will correctly be

[23]The judgments that we model here are thus examples of "implicit cognition" of the kind discussed by Cunningham (2013), Kleinberg *et al.* (2018), and Li and Camerer (2022), among others. As in Bayesian models of perceptual illusions (e.g., Feldman, 2013), our use of Bayesian calculations follows from a hypothesis that the cognitive module makes efficient use of the information available to it, not that it reflects conscious reasoning.

[24]See the references above, and additional references in Khaw *et al.* (2021).

judged larger the other (Piazza, 2010).

This fact can be captured by hypothesizing that the ANS representation of numbers are ordered on a logarithmically compressed "mental number line." According to this hypothesis, the accuracy of comparison of two numbers is an increasing function of the distance between their representations; and the distance between the representations of two numbers is proportional to the difference in their logarithms (Dehaene, 2003; Verguts and Fias, 2004). We can capture this idea in a simple model by supposing that the neural representation of a given numerical magnitude $|X|$ can be summarized by a single real number r_x, drawn from a distribution

$$r_x \;\sim\; N(\log|X|,\, \nu_x^2) \tag{2.1}$$

where the noise variance ν_x^2 is independent of the number represented. Then if we suppose that a second array will be judged on a given trial to involve a larger number of items than a first array if and only if the two noisy representations satisfy $r_2 > r_1$, the probability of this occurring is predicted to be an increasing function of $\log(X_2/X_1)$ — i..e., the percentage difference rather than the absolute difference in the numbers.

Biases in subjects' estimates of the number of items in an individual array are also well-documented. It has been noted since the work of Jevons (1871) that "there is a clear tendency to over-estimate small numbers and to under-estimate large ones" (p. 282).[25] (See Figure 6 below for examples of this bias.) The average estimated number is commonly an increasing, strictly concave function of the true number; this is often modeled by plotting estimated numerosity against true numerosity in a log-log plot (e.g., Krueger, 1984; Kramer *et al.*, 2011) and finding a slope $\gamma < 1$ for a regression line fitted to the data.

Estimation bias of this type can be understood as an efficient response to the fact that estimates are based on a noisy representation of the stimulus rather than its exact properties, if one supposes that the response rule producing an estimate $\hat{X}(r_x)$ is optimized for a particular prior distribution over possible values of X, representing the kind of numbers that one can expect to encounter in a given context.[26] In fact, a log-linear relationship between the mean observed estimate $\mathrm{E}[\hat{X}\,|X]$ and the true number X will be optimal (in the sense of minimizing the mean squared error of the estimate) if the imprecision in the noisy representation is of the kind postulated in (2.1), and the response rule is optimized for a log-normal prior,

$$\log|X| \;\sim\; N(\mu_x,\, \sigma_x^2). \tag{2.2}$$

In this case, the posterior distribution for $|X|$ conditional on r_x will also be log-normal,

[25] Economists who are skeptical of the relevance for our field of studying the accuracy of numerical judgments about visual arrays might note that William Stanley Jevons, who introduced the concept of marginal utility as a determinant of economic valuations, was well aware of, and indeed an important contributor to the experimental study of number perception.

[26] Note that the prior should be different in different contexts, and the optimal response rule $\hat{X}(r_x)$ will accordingly be different, even if one supposes that in each case the rule adapts so as to minimize the mean squared error of the estimates. In fact, adaptation of this kind of the relationship between the mean observed estimate and the true number is observed in studies of numerosity estimation; see, for example, Xiang *et al.* (2021), who compare estimation biases in three treatments in which numbers of drawn from three different ranges, each of the same width but with different midpoints.

with a mean $\hat{X}(r_x) \equiv \mathrm{E}[|X|\,|r_x]$ such that

$$\log \hat{X}(r_x) \;=\; (1-\gamma_x)\bar{\mu}_x \;+\; \gamma_x \cdot r_x, \tag{2.3}$$

where

$$\gamma_x \;\equiv\; \frac{\sigma_x^2}{\sigma_x^2+\nu^2} \;<\; 1, \tag{2.4}$$

and $\bar{\mu}_x$ is the logarithm of the prior mean of $|X|$.[27] It follows that for any true X, the estimate $\hat{X}$ will be log-normally distributed, with a mean and standard deviation that scale as

$$\mathrm{E}[\hat{X}\,|X] \;\sim\; |X|^{\gamma_x}, \qquad \text{s.d.}[\hat{X}\,|X] \;\sim\; |X|^{\gamma_x}. \tag{2.5}$$

Thus the model predicts a log-linear relationship between the true number and the mean estimate, with a slope less than 1. It also predicts that the variability of estimates will be larger in the case of larger numerosities; in particular, it predicts the "scalar variability" relationship that is found empirically, according to which the standard deviation of estimated numerosity grows in proportion to the mean estimate (Piazza, 2010).

This model of estimation bias has direct implications for the stake-size effects that should be expected in lottery valuations, if the imprecision in the mental representation of monetary payoffs is of the same kind as in perceptions of numerosity. In our experiment, the decision problem presented on a given trial is specified by two numbers, the non-zero monetary outcome X and the probability p with which it will be received. Suppose that the magnitude $|X|$ has a noisy representation r_x drawn from (2.1), and that the relative probability of the two outcomes also has a noisy representation r_p, drawn from a distribution that depends on the true value of p.[28] We assume that the sign of X is understood without error. We assume that the DM's WTP must be chosen on the basis of the noisy representation $\boldsymbol{r}$, a vector with two components (r_p, r_x).

The incentive scheme in our experiment implies an expected financial loss from inaccurate bidding proportional to $(WTP-EV)^2$, where $EV \equiv pX$ is the lottery's expected value.[29] An optimal bidding rule (one that maximizes the DM's expected financial wealth) will therefore require that

$$WTP \;=\; \mathrm{E}[pX\,|\boldsymbol{r}] \tag{2.6}$$

on each trial. If we assume both that the random variables r_p, r_x are conditionally independent and that p and $|X|$ are independent random variables under the prior,[30] then we can write $\mathrm{E}[pX\,|\boldsymbol{r}] = \hat{p}(r_p)\cdot\hat{X}(r_x)$.

Then if the imprecision in r_x is of the form (2.1) and the prior over payoff magnitudes for which the decision rule is optimized is of the form (2.2), it follows from (2.3) that (2.6) implies that

$$\log WTP \;=\; \log\hat{p}(r_p) \;+\; (1-\gamma_x)\bar{\mu}_x \;+\; \gamma_x\cdot r_x, \tag{2.7}$$

[27]See the Appendix, section D.1, for details of the calculations. The predictions of this model are compared to empirical data in Heng *et al.* (2025).

[28]This is specified more precisely in section 3.1 below.

[29]See the Appendix, section A.2, for further discussion.

[30]Note that this is a feature of the probability distribution from which the lotteries faced by each of our subjects is drawn: the same set of values for $|X|$ are faced, regardless of p and of the sign of X.

where γ_x is again given by (2.4). Hence conditional on the values of p and X on a given trial, the mean relative risk premium should be given by a relation of the form

$$\mathrm{E}[\log(WTP/EV)\,|p, X] \;=\; \alpha_p \,+\, \beta_p \log|X|, \tag{2.8}$$

where

$$\alpha_p \;=\; \mathrm{E}[\log\hat{p}(r_p)\,|p] \,-\, \log p \,+\, (1-\gamma_x)\bar{\mu}_x, \qquad \beta_p \;=\; -(1-\gamma_x). \tag{2.9}$$

Moreover, if the numbers denoting monetary payoffs are encoded with the same precision whether these represent potential gains or losses, the coefficients of the relation (1.2) are predicted to be the same regardless of the sign of X.

Because (2.4) implies that $0 < \gamma_x < 1$, (2.9) implies that $-1 < \beta_p < 0$. Thus the model predicts the sign of the stake-size effects that are observed in Figures 3-4 in both the gain and loss domains. It also explains why these stake-size effects should be log-linear; why the slope β_p should be negative, but less steep than -1; and why the affine relation (1.2) is the same in the gain and loss domains for each value of p.

However, this simple version of a cognitive noise model doesn't yet explain why stake-size effects are most pronounced when p is small. This depends on recognizing that the precision with which monetary payoffs are repesented need not be fixed.

2.2 Endogenous Precision of Number Representation

Our result (2.4) implies that γ_x depends on the degree of imprecision in number representation, measured by the noise parameter ν. Holding fixed the prior (i.e., the statistics of the numbers encountered in a given environment), increased imprecision (larger ν) implies that an optimal estimation criterion should involve a lower value of γ_x. This in turn means greater concavity of the relationship between X and $\mathrm{E}[\hat{X}\,|X]$, hence a greater degree of over-estimation of small numbers and greater under-estimation of larger ones.

Xiang *et al.* (2021) demonstrate such an effect in the case of numerosity estimation in two ways. They ask subjects to rate their subjective confidence in their numerosity estimates, and they seek to manipulate the precision of number representations by varying the time that subjects are allowed to view the image before submitting their estimates. Both in the case of trials on which below-average confidence is expressed, and trials on which subjects are allowed less time, estimates are found to exhibit greater random variability, as our model would predict if ν is larger on these trials. And in both of these cases, the authors also find reduced sensitivity of the mean estimate $\mathrm{E}[\hat{X}\,|X]$ to increases in X, as the model of optimal adaptation to cognitive noise would predict, if subjects are aware of the greater imprecision in their number representations in these cases, and able to adapt their estimation rule accordingly.[31]

There is also evidence suggesting that people are able to vary the precision of their number representations in response to changes in the incentives for greater precision. Prat-Carrabin and Woodford (2026) report a numerosity estimation experiment in which the

[31] There is evidence from many other kinds of cognitive tasks, numerical and otherwise, that people have some degree of metacognitive awareness of the degree of precision of their mental processing — what Enke and Graeber (2023) call "cognitive uncertainty." In many of these cases, greater awareness of imprecision is associated with reduced sensitivity of behavioral responses to parameter variation (Enke *et al.*, 2026).

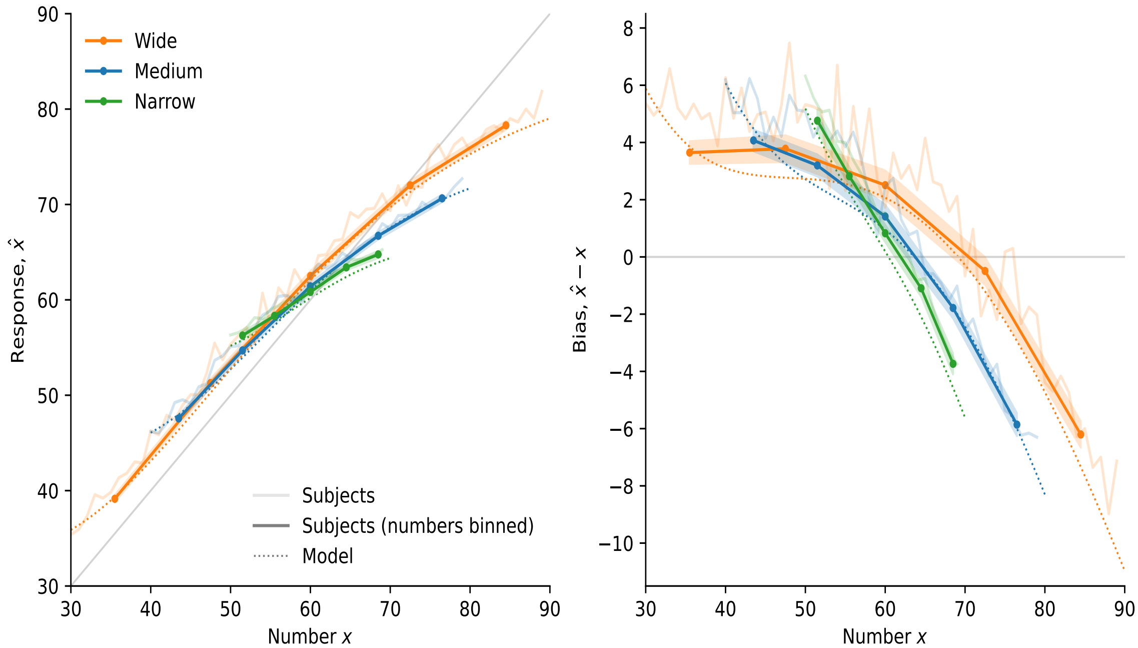


Figure 6: Bias in numerosity estimates when numbers are drawn from three different ranges (indicated by different colors). Left panel: average estimate as a function of true numerosity, for each of the three treatments. Right panel: bias as a function of true numerosity. In each panel, the faint thin lines plot the raw data; dots connected by thicker lines smooth the curves by binning the data; and dotted lines show the predictions of the endogenous precision model. (Source: Prat-Carrabin and Woodford, 2026.)

numbers presented are drawn from three different ranges, with varying widths but the same average value, in different treatments.[32] They find that the distribution of responses to a given stimulus differs depending on the context in which the visual image is presented; in particular, for a given true numerosity around the midpoint (that occurs in all three treatments), the variability of responses is smallest under the "narrow" treatment and largest under the "wide" treatment. The average estimate associated with a given true numerosity also differs across the treatments, as shown in Figure 6. In each case, smaller numbers are over-estimated while larger numbers are under-estimated; but the degree bias declines most steeply with increases in the true number when the numbers presented are drawn from only a narrow range.

This context-dependence of both the variability of estimates and their average bias can be explained as a consequence of efficient adaptation to the different priors appropriate to different contexts, if we suppose that the precision of number representation also adjusts. Prat-Carrabin and Woodford propose a model in which the noise variance in (2.1) is equal to $\nu_x^2 = \omega^2/\tau$, where ω^2 is the variance of log numerosity in a given context, and τ is a measure of the intensity of activity in the neural circuit that represents the number.[33] Precision can

[32]Numbers are drawn from three ranges: narrow (50-70), medium (40-80), and wide (30-90) ranges. See Figure 6 for the mean estimates for numerosities drawn from each of these ranges.

[33]For a fixed value of τ, scaling the noise variance in proportion to ω^2 implies a fixed channel capacity of

be increased at a cost that is proportional to τ. Hence the model postulates that in a given context, both the value of τ and the estimation rule $\hat{X}(r_x)$ adjust so as to minimize a hybrid loss function

$$\mathrm{E}[(\hat{X}(r_x) - X)^2] + c \cdot \tau, \tag{2.10}$$

where $c > 0$ is the per-unit cost of additional precision.

The model implies that in an environment where the variance of $\log X$ is greater, the mean squared error (MSE) is reduced to a greater extent by a marginal increase in τ, and hence the optimal value of τ increases, though not by so much as to completely eliminate the increase in MSE. Hence the model predicts that in an environment with larger ω, the MSE should increase, but by a factor less than proportional to ω^2, as found in the experimental data.[34] The increase in τ also implies that the steepness with which estimation bias decreases with increases in X should be smaller in a higher-ω environment. Figure 6 shows that this occurs, and indeed to almost exactly the extent predicted by the endogenous precision model (indicated by dotted lines in the figure, when a single value of c is fitted to all three treatments).

This suggests that estimation bias depends on the incentive for precision, and also suggests a particular cost function for precision.[35] We now consider the implications of assuming the same model of endogenous precision in mental representations of the monetary payoffs offered by lotteries.

2.3 Endogenous Precision and Variable Stake-Size Effects

We now consider the possibility that the noise parameter ν_x^2 in (2.1) can vary depending on the DM's representation of the probability that the lottery pays off, writing

$$r_x \sim N(\log |X|, \nu_x^2(r_p)), \tag{2.11}$$

the Gaussian channel that yields the noisy reading of number as its output (Cover and Thomas, 2001, chap. 10). Increasing τ increases the information transmission to the same extent as τ independent signals, each with variance ω^2.

[34] Prat-Carrabin and Woodford find that apart from a constant additive contribution to the MSE attributed to motor noise, the MSE grows roughly in proportion to ω (rather than ω^2) across their three treatments.

[35] Prat-Carrabin and Woodford (2026) provide further support for the model by measuring how accuracy varies with the range of numbers used in a particular experiment in the case of a different number-processing task, in which the financial loss from inaccuracy is of a different form from (2.10), so that it is no longer optimal for τ to grow roughly in proportion to ω. Here we emphasize their results for the numerosity-estimation task, because the measure of financial loss from imprecise number representation in (2.10) is more closely analogous to the losses from imprecision in our lottery-valuation task (see below), and because we are especially interested in estimation biases of the kind shown in Figure 6. But it is worth noting that their alternative task involves numbers that are presented using number symbols (Arabic numerals), like our experiment. Thus the success of the endogenous precision model in matching their alternative task suggests that it should also apply to noisy representations of symbolically presented numbers.

where the dependence of ν_x on r_p must be determined.[36] As in Prat-Carrabin and Woodford (2026), we assume a cost of more precise representation that is linear in precision,[37] so that both the function $\nu_x(r_p)$ and the response rule $C(\boldsymbol{r})$ are adapted to minimize a total loss of the form

$$\mathrm{E}[(C - pX)^2 \; + \; A \cdot \nu_x^{-2}], \tag{2.12}$$

where $A > 0$ is a cost per additional unit of precision, scaled to be in the same units as the first term, and we now use the simpler notation C for the DM's bid (WTP).

As above, the optimal bidding rule will be given by (2.6). If we simplify the analysis by supposing that p is encoded with perfect precision, then (2.12) reduces to

$$\mathrm{E}[p^2 \cdot \mathrm{E}[(\hat{X} - X)^2] + A \cdot \nu_x^{-2}(p)],$$

where the outer expectation is over the distribution of possible values of p and the inner expectation is over possible values of X and r_x. For each value of p, one would then want to choose $\nu_x(p)$ and $\hat{X}(r_x; p)$ so as to minimize the expression inside the outer expectation symbol. This is a loss function of exactly the form (2.10), except that here the relative weight on precision is A/p^2, whereas above it was $c\omega^2$. Thus reducing p should have the same effect on the optimal choice of ν_x as increasing ω in the experiment of Prat-Carrabin and Woodford (2026): it should increase ν_x, but by a factor less than the increase in $1/p$.

On the other hand, if we are interested in the consequences for the ratio ν_x/ω, then reducing p in the lottery-valuation task is like *reducing* ω in the numerosity-estimation task: the optimal ν_x/ω increases in both cases (in the latter case, because ν_x increases less than in proportion to ω). In the case of a log-normal prior (2.2) for $|X|$, (2.4) implies that the elasticity γ_x is a decreasing function of ν_x/σ_x, and hence of ν_x/ω. The endogenous precision model would accordingly predict that γ_x should be smaller for smaller p; hence (2.9) implies that the elasticity β_p should be more negative for smaller values of p, just as the slope of the bias is more negative in the case of a narrower range in Figure 6.

The discussion thus far has assumed for simplicity that p is observed with perfect precision. But our results in section 4.2 indicate that our subjects' bids are best explained under a hypothesis that they are based on a noisy representation r_p from which p cannot be exactly inferred. In this more general case, (2.12) can be expressed as the expectation over possible realizations of r_p of a conditional loss

$$\begin{aligned} L(r_p) &= \mathrm{E}[(C - pX)^2 \,|r_p] + A \cdot \nu_x^{-2}(r_p) \\ &= \mathrm{E}[(\hat{p}\hat{X} - pX)^2 \,|r_p] + A \cdot \nu_x^{-2}(r_p) \\ &= \mathrm{E}[X^2] \cdot \mathrm{var}[p\,|r_p] \; + \; \hat{p}(r_p)^2 \cdot \mathrm{E}[(\hat{X} - X)^2 \,|r_p] \; + \; A \cdot \nu_x^{-2}(r_p). \end{aligned} \tag{2.13}$$

[36] One might instead assume that the noise variance should depend on the true value of p. But then in evaluating the conditional expectation (2.6), the precision with which r_x has been encoded will not be known, and instead would have to be (imperfectly) infered from r_p. Our analytical derivations below are simplified by supposing that p is first encoded by r_p, and that the value of r_p then influences the encoding of $|X|$.

[37] The assumption of a cost of precision that is linear in precision is also sometimes used in rational inattention models in economics; see, e.g., Myatt and Wallace (2012). A similar cost function is found to account well for endogenous variation in the precision of visual working memory in van den Berg and Ma (2018). We provide a cognitive process interpretation of this cost function in the Appendix, section B.

Thus (2.12) is minimized if for each value of r_p, both $\nu_x(r_p)$ and the estimation rule $\hat{X}(r_x; r_p)$ are chosen to minimize (2.13).

Since the first term in (2.13) is independent of both ν_x and the estimation rule, we obtain again the same kind of optimization problem as was just considered, but now with p replaced by $\hat{p}(r_p)$, the posterior mean estimate of p implied by the noisy representation r_p. Thus we can conclude, using the same reasoning as above, that values of r_p that imply that p is likely to be low (i.e., that mainly occur when p is low) make it optimal for $\nu_x(r_p)$ to be larger, implying a lower value for $\gamma_x(r_p)$.

In this more general case, the optimal bidding rule is again of the form (2.7), except that now γ_x must be replaced by $\gamma_x(r_p)$, the quantity defined by (2.4) when we replace ν_x by $\nu_x(r_p)$. Averaging over the possible noisy representations $\boldsymbol{r}$ that may be produced by a given lottery $(X;p)$, we again obtain a log-linear relation of the form (2.8), except that in (2.9) we must now replace γ_x by

$$\gamma_p \equiv \mathrm{E}[\gamma_x(r_p)\,|p],$$

a quantity between 0 and 1 that depends on p. Because γ_p now depends on p, the predicted slope coefficient β_p does as well. And to the extent that lower values of p are associated (on average) with lower values of r_p and hence lower values of $\gamma_x(r_p)$, as is true in the particular model of representation of probabilities proposed below, the model will imply that β_p should be more negative for smaller values of p, as we observe in Figures 3-4.

Thus the model provides an explanation not only for the sign of stake-size effects and for their log-linearity, but also for the fact that they are more pronounced (in both the gain and loss domains) in the case of lotteries $(X;\, p)$ for which p is small. We turn now to the question whether the model can be parameterized so as to account quantitatively for the risk premia reported in Figures 3-4.

3 Endogenously Imprecise Lottery Valuation: A Quantitative Model

Here we specify additional details of the quantitative model that we fit to the bidding behavior of the average subject, discussed in section 1.

3.1 Imprecise Representation of Probabilities

In the discussion above, we have allowed for the possibility that the relative probabilities of the two lottery outcomes are represented only imprecisely, by a random variable r_p. In our baseline quantitative model, we follow authors such as Khaw *et al.* (2021), Enke and Graeber (2023), and Vieider (2024) in assuming that this quantity is independently drawn on any trial from a conditional distribution

$$r_p \sim N\big(\log\frac{p}{1-p},\, \nu_z^2\big), \tag{3.1}$$

where ν_z^2 is independent of p. The log odds transformation results in a measure of relative probability that can be located anywhere on the real line, making it possible to assume Gaussian noise in this representation.

The log odds transformation also implies that nearby probabilities are most sharply discriminated when they are either very small (near 0) or very large (near 1), and this has been found to be the case, both in the case of perceptual judgments of relative frequencies (Tong *et al.*, 2026), and judgments of the relative size of fractions or proportions, the terms of which are specified using number symbols (Frydman and Jin, 2025). Eckert *et al.* (2018) find that the discriminability of two different relative frequencies ($p_1/1-p_1$ versus $p_2/1-p_2$) is increasing in the "ratio of ratios," or equivalently, in the difference between the log odds in the two cases, just as (3.1) would predict.[38]

Of perhaps even more relevance for our application to lottery valuation, Zhang and Maloney (2012) review evidence on biases in estimation of relative frequency or proportions, and show that these biases are roughly linear in log odds — that is, the log odds implied by subjects' estimates are an affine function of the true log odds. This kind of estimation bias can easily result from use of an optimal estimation rule when the probabilities or proportions in question are imprecisely represented with noise of the kind specified in (3.1). [39]

Estimation bias of this kind provides a simple explanation for the "fourfold pattern" of risk attitudes documented by Tversky and Kahneman (1992), and seen in our data as well.[40] Even in the absence of any imprecision in the representation of monetary payoffs, (2.6) implies that an optimal bid for a lottery $(X; p)$ should equal $\hat{p}(r_p) \cdot X$, so that the median bid should be given by $w(p) \cdot X$, where

$$w(p) \;\equiv\; \mathrm{med}[\hat{p}(r_p)\,|p]. \tag{3.2}$$

The function (3.2) would then have implications similar to the probability weighting function posited in prospect theory, and if estimation bias is of the kind documented in other contexts by Zhang and Maloney (2012), the function will have the inverse-S shape required to generate the Tversky-Kahneman fourfold pattern.

Imprecise coding of probabilities of the kind specified in (3.1) also provides an explanation for the finding of Enke and Graeber (2023), that subjective uncertainty about the certainty-equivalent value of lotteries like the ones in our experiment varies as an inverse-U-shaped function of the value of p (that is, higher for intermediate values of p than for either very small or very large values). If we interpret the subjective uncertainty about lottery values in their experiment as a consequence of uncertainty about the value of p implied by a given noisy internal representation r_p,[41] then this result would also follow from (3.1), that makes nearby values of p more difficult to distinguish in the case of intermediate values of p.

Our precise predictions for the optimal bias in the estimate $\hat{p}(r_p)$ depend on the prior for which the estimation rule is optimized. In our quantitative model, we assume a prior

[38]Our hypothesis is also consistent with the suggestion by Tversky and Kahneman (1992), that people exhibit "diminishing marginal sensitivity" to information about probabilities as the probability moves farther from either of two "reference points," one at zero and the other at a probability of 1. Gonzalez and Wu (1999) provide further discussion and experimental evidence.

[39]See Appendix section C for details of the calculations.

[40]Bayesian explanations of this form for the probability weighting function of prospect theory have been offered by Fennell and Baddeley (2012), Khaw *et al.*, (2021), Enke and Graeber (2023), Vieider (2024), Bedi *et al.* (2025), and Tong *et al.* (2026).

[41]See further discussion in the Appendix, section C.2, of how our model can be used to explain the results of Enke and Graeber.

distribution for p of the form

$$\log \frac{p}{1-p} \sim \text{Uniform}\,[\mu_z - \sqrt{3}\sigma_z,\ \mu_z + \sqrt{3}\sigma_z], \tag{3.3}$$

for some parameters μ_z, σ_z, which indicate the mean and standard deviation of the prior. Our choice of this form allows the prior to be specified by two parameters, the values of which are fit to the distribution of lotteries faced in the experiment.[42]

3.2 Imprecise Response Selection

We have thus far supposed that a DM should be able to choose their expressed WTP optimally as a function of the noisy internal representation (r_p, r_x). Our quantitative model allows for a further type of cognitive noise: we also allow for imprecision in the DM's ability to recognize that a particular monetary bid C corresponds to a particular subjective sense of the value of the lottery. As with the "expression theory" of Goldstein and Einhorn (1987), our theory of how a monetary value is assigned to a lottery involves three distinct cognitive operations: (a) *encoding* of the stated lottery characteristics by an internal representation $\mathbf{r} \equiv (r_p, r_x, \text{sign}(X))$; (b) *evaluation* on the basis of the internal representation, producing a subjective sense of the value of the lottery; and (c) *expression* of that valuation using a particular response scale — here a monetary bid corresponding to a particular slider position. Our baseline model allows for imprecision in both stages (a) and (c).

We suppose that the DM associates their subjective evaluation of the lottery with a particular overt response on the basis of an imprecise internal representation of the value corresponding to each of the possible bids C. We suppose that for each potential bid magnitude $|C| > 0$, the DM has a noisy representation r_c of its value; the DM then chooses the bid (slider position) corresponding to some value $r_c = f(r_p, r_x)$ that depends on the internal representation of the lottery characteristics. (The sign of the bid is always the same as the sign of X: this is enforced by our experimental protocol, and we assume no wish by the DM to express anything else.)

While we assume a non-degenerate distribution of possible values of r_c for any given bid magnitude $|C|$, it makes sense to suppose that the DM is aware of the fact that the ordering of positions on the slider corresponds to the ordering of the values associated with those bids. A simple way of ensuring this is to suppose that

$$r_c = \log|C| + \epsilon, \qquad \epsilon \sim N(0, \nu_c^2), \tag{3.4}$$

where the scalar quantity ϵ varies randomly from trial to trial, but is the same for all $|C|$ on a given trial. Note that (3.4) is the same kind of model of imprecision in the representation of monetary amounts as we assume in the case of the lottery payoff magnitude $|X|$, and can be justified on similar grounds. This implies that on any given trial, the DM's response C is

[42] A truncated uniform distribution better fits the set of values for the odds ratio used in our experiment than a Gaussian distribution would. Of course, the distribution sampled from is not literally a uniform distribution; only a discrete set of values of p are used, as shown in Figures 3 and 4. But we assume that the DM's response rule is optimized for general features of the environmental statistics — the mean and standard deviation of the log odds ratios that are encountered — rather than for the precise discrete distribution of lotteries used in the experiment.

a monetary amount with the same sign as X and a magnitude that is an independent draw from a log-normal distribution,

$$\log|C| \sim N(f(\mathrm{r}), \nu_c^2). \tag{3.5}$$

Despite the unavoidable randomness of the response specified in (3.5), we assume that the target f is optimal conditional on the internal representation $\mathbf{r}$ of the lottery currently under consideration. (This generalizes our assumption in the discussion above that C could directly be chosen as a function of $\boldsymbol{r}$.) Hence f is chosen so as to minimize the expected loss (2.12), when the joint distributions of $p, X, \mathbf{r}$, and C are determined by prior distributions (2.2) and (3.3) and the conditional distributions specified in (2.11), (3.1), and (3.5).

The optimal bidding rule can be shown to be[43]

$$f(\mathrm{r}) \;=\; \log\hat{p}(r_p) \,+\, (1-\gamma_x(r_p))\bar{\mu}_x \,+\, \gamma_x(r_p)r_x \,-\, \frac{3}{2}\nu_c^2, \tag{3.6}$$

generalizing (2.7). In the case of response noise, the median bid is shaded downward (in absolute size) by a constant percentage that depends on the value of ν_c^2, to take account of the multiplicative error in bidding. It follows that the mean relative risk premium should still satisfy an affine function (2.8) of $\log|X|$; (3.6) simply requires that we add an additional constant term to the solution (2.9) for the intercept α_p.

Note that optimality implies that the target f (but not the DM's actual bid C) should be a deterministic function of $\mathbf{r}$. Note also that we do not, as in some models of response noise, assume that the DM chooses a target that would be optimal in the absence of such noise, even though the actual response differs from the target owing to the noise. We instead assume that the function $f(\mathrm{r})$ is optimized for the particular degree of cognitive noise to which the DM is subject — taking into account both the encoding noise in the internal representations *and* the fact that the DM's bid will involve response noise (if $\nu_c > 0$).

3.3 Declining to Bid

As noted in section 1, on a few trials subjects submit bids of \$0, which we interpret as declining to bid on that lottery. In our quantitative model of these data, we suppose that the DM's decision actually has two stages: a first decision whether to bid at all, followed by a second decision about which (non-zero) bid to make, only in the case that the first decision was to bid. We further suppose that the decision in each stage is optimized to serve the DM's overall objective, subject to the constraint that each decision must be made on the basis of an imprecise awareness of the precise decision problem that is faced on that trial. In such a two-stage analysis, one of the benefits of deciding in the first stage not to bid will be avoidance of the cognitive costs associated with having to decide what bid to make in the second stage.[44] The cognitive costs associated with undertaking a second-stage decision should include the cost of encoding (or retrieving) the magnitude of the monetary payoff with a greater degree of precision, as specified in (2.12) above, but they could include other

[43]See the Appendix, section D.2, for details of the calculation.

[44]See Khaw *et al.* (2017) for an example of a complete model of a two-stage decision of this kind, in a different context.

costs as well, that have not been specified above because they do not affect our calculation of the optimal second-stage bidding rule. One piece of evidence in support of the view that a zero bid avoids cognitive costs is our observation that subjects respond more quickly on average on the zero-bid trials.[45]

In this paper, we model only the "second-stage" problem, i.e., how the subjects bid on those trials where they choose to make a non-zero bid. This is done taking as given the probability that the DM will find themselves having to choose a non-zero bid in the case of a particular lottery $(X;\, p)$, as a consequence of the first-stage decision rule.[46] The prior distribution that is relevant for the "second-stage" problem modeled above (specified mathematically by (3.1) and (2.11)) is not the frequency distribution with which the experimenters present different lotteries $(X;\, p)$, but rather the frequency distribution with which the different lotteries become the object of a second-stage decision. This depends both on the distribution of lotteries chosen by the experimenter and on the first-stage decision rule. However, in our quantitative evaluation of the model below, we fit the parameters of the assumed prior distribution to the *empirical* frequency with which non-zero bids are made on different lotteries $(X;\, p)$, and not to the frequency distribution of lotteries chosen by the experimenters. Given this, it is not necessary for us to model the DM's first-stage decision in order to derive quantitative predictions from our model of the second-stage decision.

In the model that we fit to our data, the prior to which the average subject's decision rule is assumed to be adapted is the joint distribution from which $(p, |X|)$ are drawn *conditional on the DM having decided to bid.* Because the probability of subjects' declining to bid is not independent of the values of p and X on that trial, in our numerical analysis of the model's predictions, we estimate the values of the parameters $(\mu_x, \sigma_x, \mu_z, \sigma_z)$ for the set of lotteries on which non-zero bids are made, rather than using the parameters for the distribution from which the lotteries in the experiment were drawn.[47]

Our quantitative model is then specified by assigning numerical values to seven parameters: the three parameters A, ν_z, ν_c specifying the amount of cognitive noise, and the four parameters $\mu_x, \sigma_x, \mu_z, \sigma_z$ specifying the prior to which the DM's cognitive processes are assumed to be adapted. But as explained further below, we determine values for the parameters of the prior so as to maximize the likelihood under the prior of facing the particular lotteries that are actually presented in the experiment. Thus only the noise parameters are "free" parameters with which we can seek to fit the bidding behavior shown in Figures 3 and 4. This means that in our baseline quantitative model, we have three free parameters with which to fit the 220 data moments plotted in Figures 3 and 4.

[45]The average response time (RT) on the 175 trials on which zero bids were submitted was 2.61 seconds, while the average RT on the other trials was 3.83 seconds (nearly 50 percent longer).

[46]See the Appendix, section A.4, for further discussion.

[47]Note that it is the distribution of lotteries that may be encountered, conditional on the trial being one where the DM chooses to bid, that matters for our derivation of the optimal bidding rule. Assuming that the DM learns best-fitting parameters for a joint distribution specified by (2.2) and (3.3), rather than a more exact description of the conditional joint distribution of $|X|$ and p, amounts to assuming that the DM learns within a somewhat flexible, but mis-specified, class of models, as in Esponda and Pouzo (2016). The assumption that only a small number of parameters must be learned makes the plausibility of fairly accurate estimation of those parameters in the context of a laboratory experiment greater.

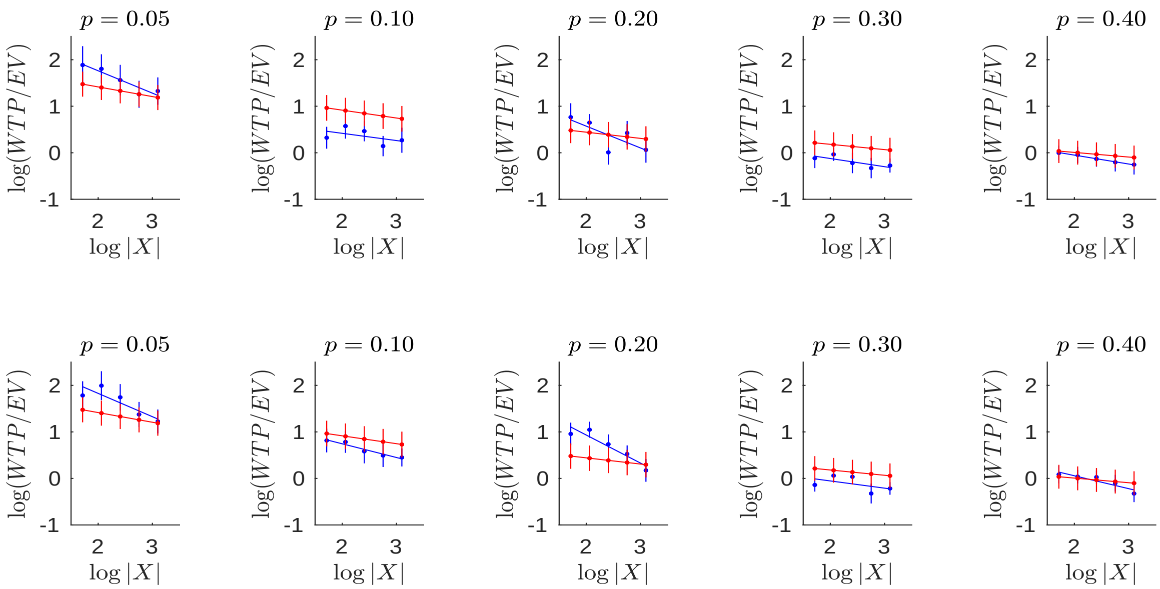


Figure 7: The same data as in Figure 3, but now compared with the predictions of the optimal bidding model with maximum-likelihood parameter estimates. (Blue: data for the average subject. Red: theoretical predictions.)

4 Assessing the Quantitative Fit of the Cognitive Noise Model

We test the complete set of predictions of the model set out above by finding the values of the three free parameters A, ν_z, and ν_c that maximize the likelihood of the data moments. As in our atheoretical modeling of the data in section 1, we express the likelihood of the experimental data as a function of the mean $m(p_j, X_j)$ and variance $v(p_j, X_j)$ of the distribution of bids for each lottery j specified by characteristics (p_j, X_j), and the number of trials N_j on which that lottery is evaluated. This amounts to approximating the predicted distribution of bids for any lottery, as a function of the model parameters, by a log-normal distribution.[48]

4.1 Fit of the Baseline Model

The theoretical data moments predicted by our model depend not only on the parameters (A, ν_z, ν_c) specifying the degree of cognitive imprecision on the part of the DM, but also on the parameters $(\mu_z, \sigma_z, \mu_x, \sigma_x)$ specifying the prior distribution over possible lotteries. Thus we estimate values for all seven parameters, so as to maximize a complete likelihood function of the data, taking into account both the likelihood of the lottery characteristics presented on the different trials (under a given parameterization of the prior) and the likelihood of the subjects' bids on those trials (given our model of noisy internal representations and optimal

[48]We assumed such a log-normal distribution (1.1) in the case of our atheoretical data characterizations. This must be regarded as only an approximation in the case of our model of optimal bidding subject to cognitive noise; see further discussion in Appendix section E.1.

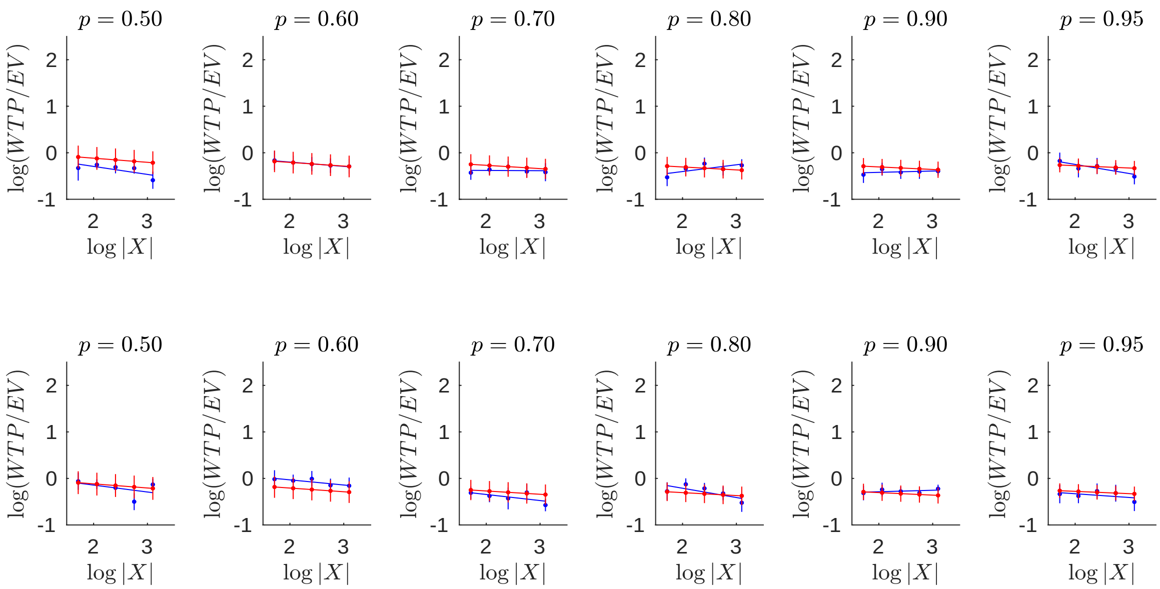


Figure 8: Continuation of Figure 7 for probabilities $p \geq 0.50$.

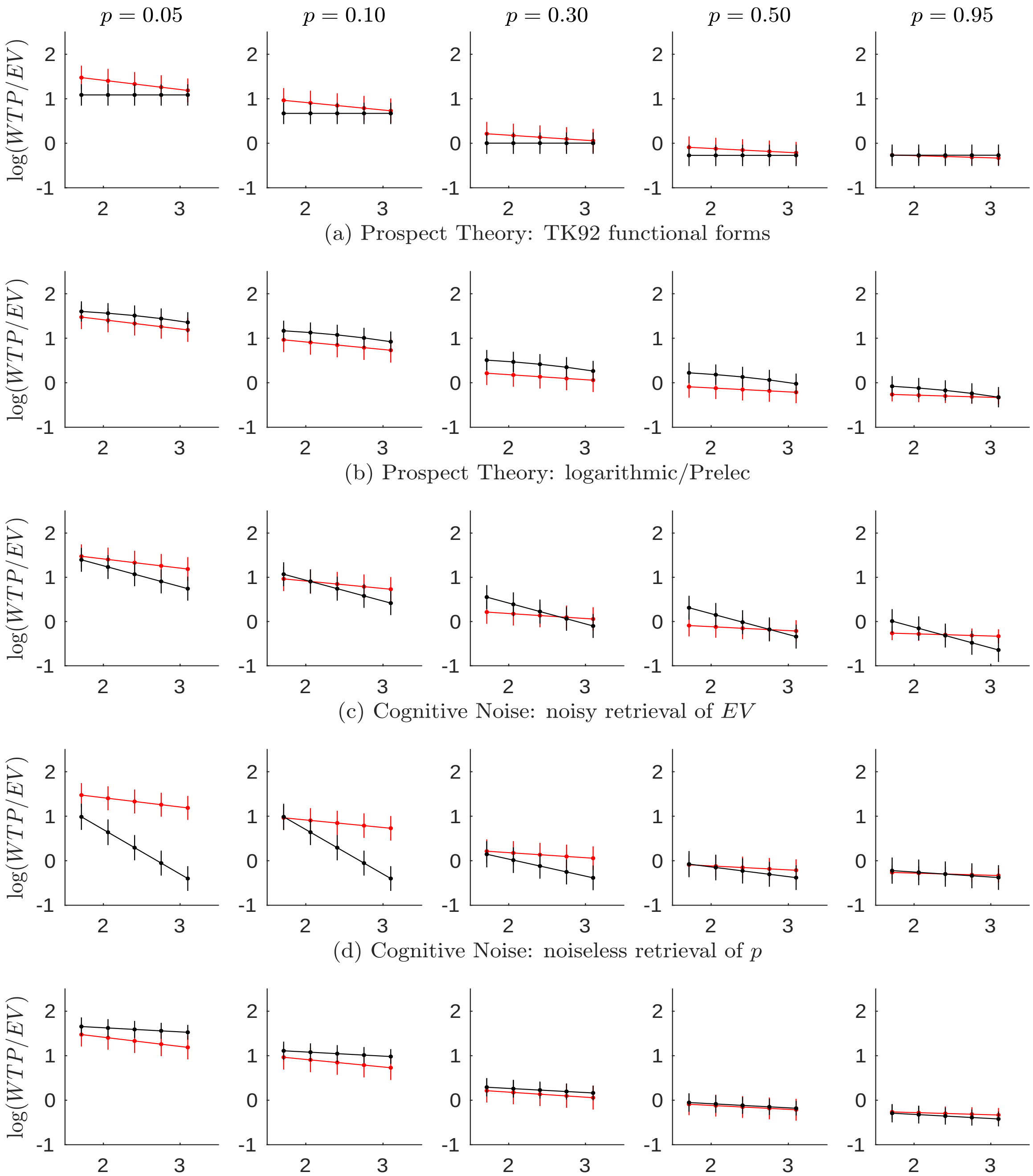
bidding).[49]

Figures 7 and 8 (presented using the same format as in Figures 3 and 4) show to what extent the predicted moments match the "average subject" moments when the parameters are chosen to maximize the (approximate) likelihood function.[50] The fit is not as good as that of the best-fitting affine model, shown in Figures 3 and 4; the maximized log-likelihood is a good deal lower, as shown on the bottom line of Table 1. However, the optimizing model has many fewer free parameters than the atheoretical affine model, and the BIC associated with the optimizing model is much lower than that of the affine model, as is also shown on the bottom line of Table 1. In fact, the BIC of the optimizing model is well below that of the best-fitting of the atheoretical models discussed above, namely the restricted version of the bounded symmetric affine model (with $\beta_p = 0$ for all $p \geq 0.3$). The Bayes factor for the optimizing model is correspondingly larger (indeed, larger by a factor greater than 10^{21}).

As an alternative benchmark for judging the degree of fit of our cognitive noise model, it is useful to compare the fit to our data of another kind of parametric model (albeit without a foundation in optimization), namely prospect theory (PT). As is well known, PT provides an explanation for the fourfold pattern of risk attitudes documented by Kahneman and Tversky, and it can be specified so as to allow for stake-size effects as well. Like our baseline model, some quantitative versions of PT involve as few as three free parameters: one to specify the degree of nonlinearity of the "value function" applied to gains or losses, one to specify the degree of nonlinearity of the "weighting function" that modifies the probabilities of the different outcomes, and one to specify the degree of random error in subjects' individual

[49]Thus our model of the data assumes that the median subject's encoding precision and decision rule have been optimally adapted to the distribution of lotteries encountered in the experiment. This assumption is consistent with the evidence of failure rapid of encoding and decoding to a different distribution of lotteries in Frydman and Jin (2022, 2025).

[50]The maximum-likelihood parameter estimates for the cognitive noise parameters are shown on the first line of Table 4 in the Appendix.



(a) Prospect Theory: TK92 functional forms

(b) Prospect Theory: logarithmic/Prelec

(c) Cognitive Noise: noisy retrieval of EV

(d) Cognitive Noise: noiseless retrieval of p

(e) Cognitive Noise: exogenous precision

Figure 9: Predictions of alternative stochastic models of lottery valuation, using the same format as in Figures 7-8 (but only for selected values of p). Each row shows the predictions of the baseline model (in red) and one alternative model (in black).

responses (Stott, 2006). In Appendix section G, we present model comparison statistics for the fit of several stochastic versions of PT to the distributions of bids of our average subject.

The variant that best fits our data involves the "logarithmic" specification of the value function advocated by authors such as Bouchouicha and Vieider (2017) as a way of matching empirically observed stake-size effects, the two-parameter probability weighting function proposed by Prelec (1998), and multiplicative response noise (for a total of four free parameters).[51] Even in this case, the BIC for the PT model is equal to 3283.5, and thus worse than the value reported in Table 1 for our baseline cognitive noise model.[52]

4.2 Comparing Alternative Models of Cognitive Noise

We have focused thus far on global measures of fit for our complete model, taken as a package. Here we consider the contribution that particular features of the baseline model make to its empirical success. We are especially interested in determining which of the several kinds of cognitive noise are most important for the model's ability to fit our data.

Table 2 reports model comparison statistics for a variety of models, in each of which subjects' bidding rules are assumed to be optimally adapted so as to minimize the objective (2.12); but the models differ in their specification of cognitive noise. The top line recalls the log likelihood and BIC statistic for our baseline model, with three types of cognitive noise and endogenous imprecision in the representation of the monetary payoff. (These numbers are the same as those reported in Table 1.) The next line instead considers an asymmetric version of the model, in which the three noise parameters are allowed to be different in the case of lotteries involving losses rather than gains. While the log likelihood is necessarily slightly higher in the case, it is not enough higher to outweigh the penalty for the additional free parameters in the asymmetric case; the BIC is higher, implying a Bayes factor of 18 in favor of the symmetry assumption in our baseline model.[53]

The "exogenous precision" model instead assumes that ν_x is a fixed parameter for all lotteries, rather than varying with r_p as in the baseline model; the numerical value of ν_x is then a parameter to be estimated (instead of the cost function parameter A). This alternative, which requires stake-size effects to be of the same size for all p (since γ_x must be independent of r_p), reduces the likelihood of the data modestly, but does not dramatically worsen the fit of the model. (Row (e) of Figure 9 illustrates the extent to which the predictions of this model remain similar to those of the baseline model, even though it fails to capture the fact that β_p is more negative for the low values of p.)

The next set of alternatives each shut off one of the types of cognitive noise in the baseline model. The model with "no payoff noise" assumes that the value of X is encoded

[51] The degree to which the predictions of this estimated version of PT are similar to those of our baseline cognitive noise model is illustrated visually on line (b) of Figure 9 below. A version of PT that uses the functional forms proposed by Tversky and Kahneman (1992) leads to less similar predictions; for example, there would be no stake-size effects, as seen on line (a) of Figure 9. That version of also fits our data less well, as shown by the model comparison statistics in Appendix section G.

[52] See Appendix section G for details. We also show that the cognitive noise model compares favorably to each of the stochastic versions of PT considered on a measure of out-of-sample fit based on five-fold cross-validation.

[53] The remaining theoretical models in the table continue to assume the same noise parameters in the case of both gains and losses.

model	#params	LL	BIC	K
baseline model	3	-1602.5	3236.3	1
asymmetric	6	-1598.1	3242.1	18
exogenous precision	3	-1604.7	3240.7	9.0
no payoff noise	2	-1608.7	3242.4	21
no probability noise	2	-1990.0	4005.0	8.3×10^{166}
no response noise	2	-1646.1	3317.2	3.7×10^{17}
noisy retrieval of EV	2	-1966.4	3957.9	4.9×10^{156}

Table 2: Model comparison statistics for alternative specifications of the cognitive noise model.

and retrieved with perfect precision; this corresponds to a limiting case of the exogenous noise model in which $\nu_x = 0$ (or of the baseline model in which $A = 0$). The model with "no probability noise" instead assumes that the value of p is encoded and retrieved with perfect precision (i.e., that $\nu_z = 0$), but still allows for noisy coding of the monetary payoff (with ν_x optimally determined as a function of p) as well as response noise. And finally, the model with "no response noise" assumes that ν_c, so that the DM's bid is optimally chosen as a function of the internal representation $\mathbf{r}$; the noisy internal representations of p and $|X|$ are specified as in the baseline model.

We find that any of these more restrictive models fits noticeably worse than the model with all three kinds of cognitive noise. Among the three, however, the assumption of noisy coding and retrieval of the payoff values is least crucial for the model's fit; while assuming perfect retrieval of the value of X reduces the likelihood of the data by more than 6 log points, once one penalizes the more flexible model for its additional free parameter, the Bayes factor[54] in favor of the baseline model relative to this alternative is only around 21. Eliminating response noise (while keeping both kinds of noise in the internal representation) reduces the likelihood (and so raises the BIC) a good deal more. Most important of all is the noisy coding of probabilities: assuming that $\nu_z = 0$ lowers the likelihood of the data to such an extent that even allowing for both noisy coding of payoffs (with a precision that depends on the value of p) and response noise, the Bayes factor in favor of the full model against this alternative is larger than 10^{166}.

The problem with this model is illustrated in row (d) of Figure 9: the model predicts that when $|X| = \bar{X}$, the prior mean value for $|X|$, the conditional mean $\mathrm{E}[\log(WTP/EV)\,|p, \bar{X}]$ should be the same for all p — the value of p affects only the slope of the line passing through that point.[55] This means that for values of X near its prior mean, such a model cannot account for the effects of changes in p on subjects' apparent risk attitude — the "fourfold pattern" of Kahneman and Tversky. Since many of the monetary payoffs used in

[54]Note that in Table 2, the value of K indicates the factor by which each model is *inferior* to the baseline model. This convention is opposite to the one used in Table 1, where K indicates the factor by which each model is *superior* to the reference model (the unrestricted model, in that table).

[55]See the Appendix, section D.4, for the derivation of this prediction. The point at which the prediction is insensitive to the value of p corresponds to a value of $\log|X|$ slightly greater than 3 on the horizontal axis of the panels in row (d) of Figure 9.

the experiment are smaller than $\bar{X}$,[56] the pattern is to some extent captured by exaggerating the degree to which β_p is negative for small p; hence the best-fitting model of this type exaggerates the stake-size effects for small p, as shown in the figure.

The final line of Table 2 considers an alternative model in which it is assumed the process of expected value computation has access to the precise values of p and X specified by the experimenter, but that the result of this computation is a noisy reading of the lottery's true EV (i.e., the quantity pX). While the sign of the EV is assumed to be recognized without error, the DM's bid is assumed to be based on a noisy semantic representation of the magnitude $|EV|$, drawn from a distribution

$$r_{ev} \sim N(\log|EV|, \nu_{ev}^2),$$

by analogy with our model (2.11) of the noisy internal representation of monetary payoffs. The DM's bid is drawn from a distribution (3.5), where the function $f(r_{ev})$ is optimally chosen to minimize the objective (2.12) as in our other models. This is a model with only two kinds of cognitive noise — noise in correctly retrieving the correct EV, and response noise given the DM's subjective sense of the lottery's value — and correspondingly two free parameters (ν_{ev} and ν_c).

The predictions of this model, illustrated in row (c) of Figure 9, fit the bidding of the average subject considerably worse than those of our baseline model. The model implies the same stake-size effects for all p; but worse, it ties the strength of stake-size effects to the size of the effect of reductions in p on the relative risk premium, resulting in both an exaggeration of the predicted stake-size effects and an underestimation of the predicted effects of changes in p. Note that we are able to strongly reject this model only because our dataset includes separate variation in both p and $|X|$; we would not be able to discriminate between our baseline model and a model of noisy EV retrieval if we considered only a set of lotteries in which the payoff size varies with no variation in p (as in Khaw *et al.*, 2021) or a set of lotteries in which the probability varies but with no variation in the monetary payoff (as in Enke and Graeber, 2023).

4.3 Dependence of Model Parameters on the Number of Trials

Thus far we have fit the parameters of our cognitive noise models to the data moments for an average subject, but the moments of the bids of each individual subject are not identical to those of the "average subject." A notable way in which the moments differ across subjects is a visible difference between the behavior of the 13 subjects who each evaluated 400 lotteries and the 15 subjects who each evaluated 640 lotteries.[57] In the Appendix, we report the outcome of an exercise in which separate versions of our baseline model are fit to the data of two different "average subjects," one representing average behavior of the 400-trial subjects and the other average behavior of the 640-trial subjects. We show (on the basis of a comparison of BIC statistics that penalize the additional parameters when allowing the two average

[56] The skewness of the log-normal prior distribution implies that the prior mean is larger than the median payoff magnitude in the experiment.

[57] See Appendix section A.3 for the set of lotteries evaluated by each of the several groups of subjects, and Appendix section F for comparison of the bidding behavior of individual subjects when classified by the number of trials that they complete.

subjects to differ) that the data are better fit by a model allowing the parameters to differ for the two average subjects than one that fits a common set of parameters to the moments of both average subjects.

Thus we can improve the fit of the model, relative to what is indicated by the fits shown in Figures 7 and 8, by allowing separate parameters for the two groups of subjects. When we do so, we find that the 640-trial average subject has a much larger cost of precision in magnitude representation (and hence less precise representations of the monetary payoffs), and noisier internal representations of the probabilities as well, though the degree of noise in the representation of the response scale is similar for both.[58]

This difference in the parameter values for the two groups of subjects may reflect greater mental fatigue or loss of concentration that one might expect in the case of the subjects who were required to complete a substantially longer series of trials. Requiring more trials appears to reduce the precision of the internal representation of both the probabilities and the monetary payoffs, but with a more dramatic effect on the representation of the monetary payoffs. Heterogeneity of this kind in our dataset is quite consistent with our interpretation of departures from risk-neutral bidding as an adaptation to cognitive noise. It is instead less obvious, in the context of a purely descriptive model such as prospect theory, why the parameters of the model should vary systematically between the subjects selected to face different numbers of and distributions of lotteries.

5 Discussion

We have shown that a model in which subjects' bids are hypothesized to be optimal — in the sense of maximizing the DM's expected financial wealth, rather than any objective that involves true preferences with regard to risk, and without introducing any free parameters representing such DM preferences — can account well for both the systematic biases and the degree of trial-to-trial variability in our subjects' data, once we introduce the hypothesis of unavoidable cognitive noise in their decision process. The model can simultaneously account for the fourfold pattern of risk attitudes predicted by prospect theory (Tversky and Kahneman, 1992), relating to the effects of varying payoff probabilities and the sign of the payoffs, and the alternative fourfold pattern of Markowitz (1952) and Hershey and Schoemaker (1980), relating to the effects of varying payoff magnitudes and the sign of the payoffs. The effects on the sign of the relative risk premium of varying the terms of a simple lottery along each of these three dimensions can be explained by a single theory, which attributes departures from risk neutrality in either direction to the way in which bids should be shaded in order to take account of cognitive noise.

Thus our model interprets systematic departures from risk-neutral valuations — despite the fact that stakes are relatively small in our experiment, so that one ought not expect subjects' marginal utility of wealth to vary much as a result of the outcome of the experiment — as a consequences of imprecision in experimental subjects' mental representations of the information presented to them, as can be done for many kinds of perceptual biases (e.g., Petzschner *et al.*, 2015; Wei and Stocker, 2015, 2017; Hahn and Wei, 2024).

[58]See the parameter estimates reported in Table 5 in Appendix section F.2.

A similar hypothesis has been used to explain departures from risk-neutrality, not only lottery-valuation tasks like the one considered here (or in papers like Enke and Graeber, 2023, or Oprea, 2024), but also in binary choices between lotteries[59] This interpretation of measured risk attitudes is consistent with a larger literature in which behavioral anomalies that have often been treated as reflecting non-standard preferences or sub-optimal heuristics are instead attributed to optimal adaptation of decision rules to the presence of cognitive noise.[60]

The explanation that we offer for departures from risk neutrality in laboratory experiments is not fundamentally different from the one given by prospect theory. Kahneman and Tversky themselves argue for the psychological plausibility of the nonlinear transformations of objective payoffs and probabilities posited in prospect theory by pointing to analogies between these nonlinearities and similar distortions in perceptual domains.[61] We regard ourselves as further developing this analogy by relating the principle of "diminishing sensitivity" referred to Kahneman and Tversky to the presence of cognitive noise, as has been done in the literature on errors in sensory judgments.

It might be wondered, however, if the deeper level of explanation offered by a cognitive noise account adds anything to the kind of descriptive account offered by prospect theory and related approaches. Here we summarize some reasons for interest in a cognitive noise account, quite apart from the ability to fit the experimental data set discussed above.

5.1 Dependence of Stake-Size Effects on the Method of Preference Elicitation

The risk attitudes of experimental subjects can also be measured by observing which lottery they choose when presented with two alternatives, rather than (as here) presenting a single lottery and eliciting a valuation for it. But it has been known for some time that the two elicitation procedures do not yield consistent measures of risk attitudes, even within the same subject.[62] This inconsistency poses a puzzle for prospect theory, or any theory that assigns values to all risky and certain prospects independent of the choice set in which they appear. Context-dependent valuations are instead possible under a cognitive noise model, given that the optimal inference from noisy representations of lottery characteristics will depend on the prior for which the decision rule is optimized, and this may reasonably depend on the nature of the choice problem.

The stake-size effects documented above provide an example of a difference between the valuations obtained by our WTP elicitation procedure and those implicit in binary choices.

[59] See, for example, Barretto-García *et al.* (2023), Bouchouicha *et al.* (2025a, 2025b), De Hollander *et al.* (2025a), Enke and Shubatt (2026), Frydman and Jin (2022, 2025), Khaw *et al.* (2021), Netzer *et al.* (2025), Steiner and Stewart (2016), Vieider (2024) and Woodford (2012).

[60] See, for example, Augenblick *et al.* (2025), Azeredo da Silveira *et al.* (2024), Bhui and Xiang (2025), Charles *et al.* (2024), Enke *et al.* (2025, 2026), Gabaix and Laibson (2022), Gershman and Bhui (2020), Natenzon (2019), Vieider (2025) and Woodford (2003, 2020).

[61] See Wakker (2025) for an enlightening discussion of the origins of prospect theory in the application of well-established psychophysical principles to the domain of judgments of risk.

[62] See, for example, Lichtenstein and Slovic (1971), Goldstein and Einhorn (1987), Harbaugh *et al.* (2010), Loomes and Progrebna (2014), Zhou and Hey (2018), Freeman and Mayraz (2019), Friedman *et al.* (2022), and Bouchouicha *et al.* (2025b).

In Khaw *et al.,* (2021) we examine choices between a simple lottery $(X; p)$ and a certain amount C, and find that they are scale-invariant: the probability of choosing the risky lottery (for a fixed value of p) is the same function of X/C for any of a range of values for C. This might seem inconsistent with the behavior of the subjects in the experiment reported here. But in Khaw *et al.* (2021) we show that scale-invariant binary choice (as long as the monetary amounts involved are all small) is an implication of the same model of logarithmic coding of payoff magnitudes as is used in section 2.1 above.

The difference is that our previous paper dealt with binary choices between a lottery and a particular certain amount, both of which varied from trial to trial, while the experiment discussed here involves a presentation of a single lottery on each trial, on which the subject bids using a scale of potential responses that is the same on all trials. In the binary-choice case, we may again suppose that the characteristics of the lottery have a noisy internal representation (r_p, r_x), drawn from the same kind of distributions conditional on the true characteristics as are posited above.[63] But the stated value of C on a given trial also needs to be represented, and it is reasonable to suppose that this quantity is represented imprecisely, in the same way as the potential payoff X is.

Letting the mental representation of the monetary amount C be summarized by a real number r_c, the DM's decision rule will maximize expected financial wealth if and only if it results in choice of the lottery if and only if

$$\mathrm{E}[pX\,|r_p, r_x] \;>\; \mathrm{E}[C\,|r_c]. \tag{5.1}$$

We have shown above that (assuming a log-normal prior for the values of $|X|$), the left-hand side of (5.1) is an increasing log-linear function of $|X|$, with an elasticity that depends on the variance ratio $\nu_x^2(r_p)/\sigma_x^2$. If we assume a similar model of the encoding of the amount C ($r_c \sim N(\log|C|,\, \nu_c^2)$), and a similar log-normal prior for C ($\log|C| \sim N(\mu_c, \sigma_c^2)$,) then the right-hand side of (5.1) will be a increasing log-linear function of $|C|$, with an elasticity that depends on the variance ratio ν_c^2/σ_c^2.

In the model of Khaw *et al.* (2021), the encoding precision parameters ν_x and ν_c are assumed to be the same, as are the prior uncertainty parameters σ_x and σ_c. As a result, the scale elasticities of the two sides of (5.1) are the same, and the condition involves only the ratio X/C, resulting in a prediction of scale-invariant choice behavior. In the model proposed above for the lottery-valuation task, instead, the DM's optimal bid is proportional to $\mathrm{E}[pX\,|r_p, r_x]$. The geometric mean of this quantity, conditional on the true quantities $(X; p)$, will an increasing log-linear function of $|X|$, but with an elasticity γ_p that is necessarily less than 1; this results in the stake-size effects discussed earlier.

The difference between the two models is not in whether the DM's cognitive process is assumed to have access to a precise reading of the value of $|C|$; both models can be specified

[63]In the model proposed in Khaw *et al.,* the probability p is treated as known with certainty (that is, the decision rule is assumed to be optimally adapted to the true value of p), and the precision parameter ν_x is assumed to be a constant, rather than being endogenously determined as in the model proposed in section 2.3. But in the experiment reported there, the value of p was the same on all trials. Thus one might assume adaptation to this particular value of p, just as we have assumed in this paper adaptation to the statistics of the distribution of values of p used in the experiment. In the absence of any variation in r_p, the model of endogenous precision proposed in section 2.3 reduces to a constant value of ν_x, and we may as well treat ν_x rather than A as the noise parameter to be fit to subjects' behavior.

to assume logarithmic encoding noise, as in (3.4), and the value of σ_c might be assumed to be the same in decision problems of the two types. The difference is in the nature of the prior that determines the optimal response to a given noisy reading of the value of $|C|$. In the binary-choice task, the DM is presented with a single value of C to consider on a given trial, and has the same prior about what $|C|$ might be on each trial: both when an unusually large value of C is offered, and when an unusually small one is. Instead, in the lottery-valuation task, the DM considers the same range of possible bids on each trial; and because the entire range is always considered, it should always be recognized that \$25 would be a relatively high bid (even if the DM has only a fuzzy sense of exactly how much \$25 is). In the binary-choice case, instead, on trials where \$25 is offered as the certain amount, an optimal decision rule treats C as likely to be worth less than that, simply because the certain amount offered on most trials is smaller.

It thus follows rather naturally from the logic of the cognitive-noise explanation for apparent risk attitudes that the "certainty equivalent" of a given risky lottery need not be the same when this is inferred from the value of C required for exactly a 50 percent probability of choosing the lottery in a binary choice, as when it is inferred from a subject's average bid in a WTP elicitation task of the kind that we study here. In this respect, cognitive noise models are not simply a proposed interpretation of the regularities predicted by prospect theory; they also predict phenomena (such as dependence of apparent risk attitudes on the method of elicitation) that prospect theory does not.

5.2 Additional Sources of Variation in Cognitive Imprecision

A notable feature of the cognitive noise model proposed above is that it implies that the degree of bias away from risk-neutral valuations should depend on the severity of the cognitive noise. We have applied this idea in our explanation for the stronger stake-size effects when the value of p is smaller: a lower probability of receiving the non-zero outcome leads to less precise representation of the magnitude of $|X|$, and hence to a stronger attenuation bias in the response of the DM's bids to variations in $|X|$. But there are a variety of other reasons why the degree of cognitive imprecision can vary, and these are often found to be correlated with the strength of observed departures from risk-neutral choices.

For example, there are differences across people in the precision of their ANS representations, and under the theory proposed here, one should expect there to be corresponding differences in both the noisiness of subjects' valuations of lotteries and in their bias away from risk-neutrality. And in fact,. Khaw *et al.* (2021) show that across their subjects, there is a significant positive correlation between an estimated subject-level index of risk aversion in binary choices between lotteries and a subject-level index of the stochasticity of choice, exactly as the cognitive noise model would predict if the difference in subjects' choice behavior is due mainly to subject-level differences in the amount of cognitive noise; see Barretto-García *et al.* (2023) for a replication.

Another possible indicator of the degree of cognitive noise, that does not require repeated presentations of the same decision problem, is a subject's reported degree of uncertainty about the correct response to give on a single trial. Enke and Graeber (2023) show that prospect-theoretic deviations from risk-neutral valuation (in each of the four quadrants of the Kahneman-Tversky "fourfold pattern") are stronger for those subjects who express greater uncertainty. As discussed further in the Appendix, section C.2, the association that they find between subjective uncertainty and the size of the departure from risk-neutral valuation is the one that our model would predict, if subjects differ in the value of ν_z and have some degree of access to the size of their personal noise parameter.

Oprea (2024) similarly finds that subjects' degree of departure from risk-neutral valuation is positively correlated with their self-reported degree of inattention to the data on payoffs and probabilities that they have been shown, and their self-assessment of the degree to which they have "guessed" rather than making a "precise (exact) decision." It is also negatively correlated with the average time that a subject takes to respond, which can be taken as an indicator of the amount of cognitive effort expended on ensuring accuracy (Rubinstein, 2013);[64] and positively correlated with the number of errors that a subject makes on a cognitive reflection test, which can also be taken as a measure of cognitive imprecision. Note that under an interpretation of prospect-theoretic biases as simply reflecting non-standard preferences, there would be no reason to expect any of these correlations with indicators of cognitive imprecision.

Another telling source of evidence that apparent risk preferences may reflect cognitive imprecision is the observation that they are unstable, changing with experience and feedback. According to the "discovered preference hypothesis" of Plott (1996), subjects in laboratory experiments can only be expected to consistently choose in accordance with their actual, considered preferences after sufficient experience with the form of task used in the experiment, involving feedback as to the consequences of their choices; and indeed, a number of authors have found that the apparent risk preferences that are expressed in relatively novel choice problems are unstable in this way (the so-called "description-experience gap": Hertwig and Erev, 2009). Perhaps the most celebrated finding of this kind is the observation that the tendency of subjects to overweight small-probability extreme outcomes, as predicted by prospect theory, occurs when lotteries involving such extreme outcomes are described to subjects (as in the experiment of Tversky and Kahneman, 1992), but disappears when subjects learn about the outcomes from repeated choices with feedback as to the consequences of choosing the lottery or not on each occasion (Hertwig *et al.*, 2004).[65] Many other authors have subsequently found that sufficient experience and feedback greatly reduce measured departures from risk-neutrality.[66]

The studies just mentioned show that it is possible to reduce the imprecision of internal representations through proper experimental design; but it is equally possible to design experimental treatments that should predictably *increase* cognitive imprecision, and this possibility is of particular interest as a test of our theory. For example, Enke and Graeber

[64]In some cases, however, a longer average response time is taken to indicate a more difficult decision, and is expected to be associated with more random choices (e.g., Alós-Ferrer *et al.*, 2021). The correlation found by Oprea (2024) can be interpreted as indicating that the main reason for differences in subjects' average response times is subject-level differences in the amount of concern for precision, rather than differences in the intrinsic difficulty of the problems faced by different subjects.

[65]Hertwig *et al.* (2004) find that when subjects must learn about the distribution of possible outcomes purely from experience, the typical result is *under-weighting* of low-probability extreme outcomes, rather than perfectly risk-neutral choice. This bias can be explained, however, as reflecting the fact that in a small sample of experience the low-probability extreme outcome will often happen not to have been observed.

[66]See, for example, Van de Kuilen and Wakker (2006); Van de Kuilen (2009); Ert and Haruvy (2017); Ert and Haruvy (2017); Charness *et al.* (2023); and Oprea and Vieider (2026).

(2023) show that increasing the complexity of the way in which information about the probability p is presented to their subjects increases the strength of prospect-theoretic biases in subjects' lottery valuations, in exactly the way that our model would imply in the case of an increase in the imprecision of the internal representation of probabilities.[67] Subjects have similarly been shown to depart farther from risk-neutral choice as a result of increased time pressure (Choi *et al.*, 2022; Kirchler *et al.*, 2017, Young *et al.*, 2012), increased cognitive load (Benjamin *et al.*, 2013; Deck and Jahedi, 2015; Gerhardt *et al.*, 2016), or acute stress (Porcelli and Delgado, 2009). These are all plausibly conditions that can be expected to reduce the precision of mental processing, perhaps in ways that can be captured by an increase in the noise parameters in our model;[68] if so, the model implies that we should expect larger departures from risk-neutrality in exactly the directions that are observed.

5.3 Why the Cognitive Noise Hypothesis Matters

In many respects, the predictions of our model are similar to those of prospect theory. As noted in section 3.1, our model of noisy representation of probability implies that the average value of $\hat{p}(r_p)$ associated with a given probability p will be an inverse-S shaped function of p, like the probability weighting function posited in prospect theory. And similarly, our model of noisy representation of the magnitude of monetary payoffs implies that the average value of $\hat{X}(r_x)$ associated with a given value of $|X|$ will be an increasing, concave function of $|X|$. When gains and losses are each encoded in this way, the average estimate of X will be an increasing, concave function of X when $X > 0$, but an increasing, convex function when $X < 0$, like the value function posited in prospect theory.[69]

But even when the theories make similar positive predictions, there are several potential advantages to deriving insensitivity to objective lottery characteristics from the nature of optimal decision making in the presence of cognitive noise, rather than simply positing nonlinear transducers as is done in prospect theory.

First, interpreting the nonlinear distortions posited by prospect theory as consequences of optimal adaptation to cognitive noise helps to explain the observed instability of prospect-theoretic parameters across settings, as reviewed above. Our interpretation offers the prospect of a more general theory that can explain when and to what degree one should expect prospect-theoretic parameters measured under one condition to predict behavior under other conditions. In this way, one can hope to use prospect theory more accurately as a predictive tool, somewhat in the spirit of the "Lucas critique" (Lucas, 1976; Sargent, 1987) of the naive use of econometric relationships for policy evaluation.

And second, our interpretation cautions against the use of estimated prospect-theoretic "preferences" as a basis for welfare evaluation. To the extent that the parameters of prospect-

[67]See the Appendix, section C.2. In a different type of decision problem, Charles *et al.* (2024) also find that the strength of a cognitive bias (imperfect pass-through from beliefs to actions) can be changed by manipulating the complexity of the information upon which subjects' beliefs should be based.

[68]Our finding in this paper that both prospect-theoretic valuation biases and stake-size effects are stronger for the subjects who completed a larger number of trials (see the Appendix, section F) suggests that fatigue may have a similar effect.

[69]An important difference, however, is that in our model, the concavity of the relationship between $|X|$ and the average estimate $\hat{X}(r_x)$ varies with the probability of the lottery paying off, owing to the dependence of $\nu_x(r_p)$ on the representation r_p.

theoretic transducers are found to change with experience and feedback, as discussed above, one should doubt that the apparent preferences estimated in situations where there has been little experience or feedback really represent considered preferences. Our analysis provides further grounds for such skepticism by showing that, even when behavior is only observed in novel situations where feedback is not given, there is good reason to regard the anomalous patterns of behavior summarized by prospect theory as reflecting errors due to cognitive noise. And our theory provides an alternative basis for welfare judgments, even if subjects' choices under ideal conditions of extensive experience and feedback cannot be observed. To the extent that one can show that peoples' choices appear to have been optimized to achieve a particular objective (here, the objective of maximization of expected financial wealth), one might plausibly regard that objective as reflecting their "true" preferences. This would still provide a basis for welfare judgments that is individualistic, in the sense that what is good for a person is inferred from what their observed behavior apparently aims to achieve, rather than reflecting what someone else wants for them (Woodford, 2018).

These cautions about the uses of prospect theory in policy design hardly imply that the departures from normative behavior documented by authors like Kahneman and Tversky are of no importance for policy analysis. The claim that people would behave in closer conformity with normative decision theory with sufficient experience and feedback should no more justify assuming that one can always assume such behavior when designing policies than the assertion that wages and prices will adjust "in the long run" so as to make real quantities independent of monetary variables would justify indifference to the extent to which different monetary policies stabilize aggregate nominal spending. We expect that the design of economic policies can be improved by taking into account the ways in which people are prone to misunderstand the circumstances under which they act. But the realization of this promise will require further progress in understanding the nature of cognitive noise and the way in which people adapt their behavior to deal with it.

References

[1] Alós-Ferrer, Carlos, Ernst Fehr, and Nick Netzer, "Time Will Tell: Recovering Preferences When Choices Are Noisy," *Journal of Political Economy* 129: 1828-1877 (2021).

[2] Anobile, Giovanni, Guido Marco Cicchini, and David C. Burr, "Number as a Primary Perceptual Attribute: A Review," *Perception* 45: 5-31 (2016).

[3] Augenblick, Ned, Eben Lazarus, and Michael Thaler, "Overinference from Weak Signals and Underinference from Strong Signals," *Quarterly Journal of Economics* 140: 335-401 (2025).

[4] Azeredo da Silveira, Rava, Yeji Sung, and Michael Woodford, "Optimally Imprecise Memory and Biased Forecasts," *American Economic Review* 114: 3075-3118 (2024).

[5] Barretto-García, Miguel, Gilles de Hollander, Marcus Grueschow, Rafael Polania, Michael Woodford, and Christian C. Ruff, "Individual Risk Attitudes Arise from Noise in Neurocognitive Magnitude Representations," *Nature Human Behaviour* 7: 1551-1567 (2023).

[6] Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak, "Measuring Utility By a Single-Response Sequential Method," *Behavioral Science* 9: 226-232 (1964).

[7] Bedi, Saurabh, Gilles de Hollander, and Christian Ruff, "Probability Weighting Arises from Boundary Repulsions of Cognitive Noise," preprint posted on *bioRXiv*, September 11, 2025.

[8] Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro, "Who Is 'Behavioral'? Cognitive Ability and Anomalous Preferences," *Journal of the European Economic Association* 11: 1231-1255 (2013).

[9] Bhui, Rahul, and Yang Xiang, "A Rational Account of the Repulsion Effect," *PsyArXiv* preprint, revision posted June 10, 2025.

[10] Bouchouicha, Ranoua, Yuchi Li, and Ferdinand M. Vieider, "Loss-Sensitivity versus Loss-Aversion," working paper, RISL$\alpha\beta$, Ghent University, February 2025a.

[11] Bouchouicha, Ranoua, Ryan Oprea, Ferdinand M. Vieider, and Jilong Wu, "Is Prospect Theory Really a Theory of Choice?," working paper, RISL$\alpha\beta$, Ghent University, September 2025b.

[12] Bouchouicha, Ranoua, and Ferdinand M. Vieider, "Accommodating Stake Effects Under Prospect Theory," *Journal of Risk and Uncertainty* 55: 1-28 (2017).

[13] Burnham, Kenneth P., and Anderson, David R., *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach,* 2d. ed., New York: Springer, 2002.

[14] Charles, Constantin, Cary Frydman, and Mete Kilic, "Insensitive Investors," *Journal of Finance* 79: 2473-2503 (2024).

[15] Charness, Gary, Nir Chemaya, and Dario Trujano-Ochoa, "Learning Your Own Risk Preferences," *Journal of Risk and Uncertainty* 67: 1-19 (2023).

[16] Choi, Syngjoo, Jeongbin Kim, Eungik Lee, and Jungmin Lee, "Probability Weighting and Cognitive Ability," *Management Science* 68: 5201-5215 (2022).

[17] Cover, Thomas M., and Joy A. Thomas, *Elements of Information Theory,* New York: Wiley, 2001.

[18] Cunningham, Thomas, "Biases and Implicit Knowledge," MPRA Paper no. 50292, posted September 30, 2013.

[19] Daunizeau, Jean, "Semi-Analytical Approximations to Statistical Moments of Sigmoid and Softmax Mappings of Normal Variables," preprint posted on *arXiv,* March 3, 2017.

[20] Deck, Cary, and Salar Jahedi, "The Effect of Cognitive Load on Economic Decision Making: A Survey and New Experiments," *European Economic Review* 78: 97-119 (2015).

[21] Dehaene, Stanislaus, "The Neural Basis of the Weber-Fechner Law: A Logarithmic Mental Number Line," *Trends in Cognitive Sciences* 7: 145-147 (2003).

[22] Dehaene, Stanislas, *The Number Sense,* revised and updated edition, Oxford: Oxford University Press, 2011.

[23] Dehaene, Stanislas, Emmanuel Dupoux, and Jacques Mehler, "Is Numerical Comparison Digital? Analogical and Symbolic Effects in Two-Digit Number Comparison," *J. Experimental Psychology: Human Perception and Performance* 16: 626-641 (1990).

[24] Dehaene, Stanislas, and J. Frederico Marques, "Cognitive Euroscience: Scalar Variability in Price Estimation and the Cognitive Consequences of Switching to the Euro," *Quarterly Journal of Experimental Psychology A* 55: 705-731 (2002).

[25] De Hollander, Gilles, Marcus Grueschow, Franciszek Hennel, and Christian C. Ruff, "Rapid Changes in Risk Preferences Originate From Bayesian Inference on Parietal Magnitude Representations," preprint, *bioRXiv,* posted October 22, 2025. [De Hollander *et al.,* 2025a.]

[26] De Hollander, Gilles, Marus Moisa, and Christian C. Ruff, "Risk Preferences Causally Rely on Parietal Magnitude Representations: Evidence from Combined TMS-fMRI," preprint, *bioRXiv,* posted January 13, 2025. [De Hollander *et al.,* 2025b.]

[27] Eckert, Johanna, Josep Call, Jonas Hermes, Esther Herrmann, and Hannes Rakoczy, "Intuitive Statistical Inferences in Chimpanzees and Humans Follow Weber's Law," *Cognition* 180: 99-107 (2018).

[28] Enke, Benjamin, and Thomas Graeber, "Cognitive Uncertainty," *Quarterly Journal of Economics* 138: 2021-2067 (2023).

[29] Enke, Benjamin, Thomas Graeber, and Ryan Oprea, "Complexity and Time," *J. European Economic Association* 23: 1838-1867 (2025).

[30] Enke, Benjamin, Thomas Graeber, Ryan Oprea, and Jeffrey Yang, "Behavioral Attenuation," working paper, Harvard University, April 2026.

[31] Enke, Benjamin, and Cassidy Shubatt, "Quantifying Lottery Choice Complexity," working paper, Harvard University, January 2026.

[32] Ert, Eyal, and Ernan Haruvy, "Revisiting Risk Aversion: Can Risk Preferences Change with Experience?" *Economics Letters* 151(C): 91-95 (2017).

[33] Esponda, Ignacio, and Demian Pouzo, "Berk-Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models," *Econometrica* 84: 1093-1130 (2016).

[34] Fehr-Duda, Helga, Adrian Bruhin, Thomas Epper, and Renate Schubert, " Rationality on the Rise: Why Relative Risk Aversion Increases with Stake Size," *Journal of Risk and Uncertainty* 40: 147-180 (2010).

[35] Feldman, Jacob, "Bayesian Models of Perceptual Organization," in J. Wagemans, ed., *Oxford Handbook of Perceptual Organization,* Oxford University Press, 2013.

[36] Fennell, John, and Roland Baddeley, "Uncertainty Plus Prior Equals Rational Bias: An Intuitive Bayesian Probability Weighting Function," *Psychological Review* 119: 878-887 (2012).

[37] Freeman, David J., and Guy Mayraz, "Why Choice Lists Increase Risk Taking," *Experimental Economics* 22: 131-154 (2019).

[38] Friedman, Daniel, Sameh Habib, Duncan James, and Brett Williams, "Varieties of Risk Preference Elicitation," *Games and Economic Behavior* 133: 58-76 (2022).

[39] Frydman, Cary, and Lawrence J. Jin, "Efficient Coding and Risky Choice," *Quarterly Journal of Economics* 137: 161-213 (2022).

[40] Frydman, Cary, and Lawrence J. Jin, "On the Source and Instability of Probability Weighting," working paper, University of Southern California, revised April 2025.

[41] Gabaix, Xavier, and David Laibson, "Myopia and Discounting," NBER Working Paper no. 23254, revised May 2022.

[42] Gerhardt, Holger, Guido P. Biele, Hauke R. Heekeren, and Harald Uhlig, "Cognitive Load Increases Risk Aversion," SFB 649 Discussion Paper no. 2016-011, Humboldt University Berlin, March 2016.

[43] Gershman, Samuel J., and Rahul Bhui, "Rationally Inattentive Intertemporal Choice," *Nature Communications* 11: 3365 (2020).

[44] Gold, Joshua I., and Hauke R. Heekeren, "Chapter 19: Neural Mechanisms for Perceptual Decision Making," in P.W. Glimcher and E. Fehr, eds., *Neuroeconomics: Decision Making and the Brain,* 2d edition, London: Academic Press, 2014.

[45] Goldstein, William M., and Hillel J. Einhorn, "Expression Theory and Preference Reversal Phenomena," *Psychological Review* 94: 236-254 (1987).

[46] Gonzalez, Richard, and George Wu, "On the Shape of the Probability Weighting Function," *Cognitive Psychology* 38: 129-166 (1999).

[47] Gonzalez, Richard, and George Wu, "Composition Rules in Original and Cumulative Prospect Theory," *Theory and Decision* 92: 647-675 (2022).

[48] Hahn, Michael, and Xue-Xin Wei, "A Unifying Theory Explains Seemingly Contradictory Biases in Perceptual Estimation," *Nature Neuroscience* 27: 793-804 (2024).

[49] Harbaugh, William T., Kate Krause, and Lise Vesterlund, "The Fourfold Pattern of Risk Attitudes in Choice and Pricing Tasks," *Economic Journal* 120: 595-611 (2010).

[50] Heng, Joseph A., Michael Woodford, and Rafael Polania, "Efficient Numerosity Estimation Under Limited Time," *PLoS Computational Biology* 21(3): e1012790 (2025).

[51] Hershey, John C., and Paul J. Schoemaker, "Prospect Theory's Reflection Hypothesis: A Critical Examination," *Organizational Behavior and Human Performance* 25: 395-418 (1980).

[52] Hertwig, Ralph, and Ido Erev, "The Description-Experience Gap in Risky Choice," *Trends in Cognitive Sciences* 13: 517-523 (2009).

[53] Hertwig, Ralph, Greg Barron, Elke U. Weber, and Ido Erev, "Decisions from Experience and the Effect of Rare Events in Risky Choice," *Psychological Science* 15: 534-539 (2004).

[54] Holt, Charles A., and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review* 92: 1644-1655 (2002).

[55] Holt, Charles A., and Susan K. Laury, "Risk Aversion and Incentive Effects: New Data without Order Effects," *American Economic Review* 95: 902-904 (2005).

[56] Jevons, W. Stanley, "The Power of Numerical Discrimination," *Nature* 3: 281-282 (1871).

[57] Kachelmeier, Steven J., and Mohamed Shehata, "Examining Risk Preferences under High Monetary Incentives: Experimental Evidence from the People's Republic of China," *American Economic Review* 82: 1120-1141 (1992).

[58] Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica* 47: 263-291 (1979).

[59] Khaw, Mel W., Ziang Li, and Michael Woodford, "Cognitive Imprecision and Small-Stakes Risk Aversion," *Review of Economic Studies* 88: 1979-2013 (2021).

[60] Khaw, Mel W., Luminita L. Stevens, and Michael Woodford, "Discrete Adjustment to a Changing Environment: Experimental Evidence," *Journal of Monetary Economics* 91: 88-103 (2017).

[61] Kirchler, Michael, David Andersson, Caroline Bonn, Magnus Johannesson, Erik Ø. Sørenson, Matthias Stefan, Gustav Tinghög, and Daniel Västfjäll, "The Effect of Fast and Slow Decisions on Risk Taking," *Journal of Risk and Uncertainty* 54: 37-59 (2017).

[62] Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Cass Sunstein, "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* 10: 113-174 (2018).

[63] Kramer, Peter, Maria Grazia De Bono, and Marco Zorzi, "Numerosity Estimation in Visual Stimuli in the Absence of Luminance-Based Cues," *PLoS ONE* 6(2): e17378 (2011).

[64] Krueger, Lester E., "Perceived Numerosity: A Comparison of Magnitude Production, Magnitude Estimation, and Discrimination Judgments," *Perception and Psychophysics* 35: 536-542 (1984).

[65] Li, Xiaomin, and Colin F. Camerer, "Predictable Effects of Visual Salience in Experimental Decisions and Games," *Quarterly Journal of Economics* 137: 1849-1900 (2022).

[66] Lichtenstein, Sarah, and Paul Slovic, "Reversal of Preferences Between Bids and Choices in Gambling Decisions," *Journal of Experimental Psychology* 89: 46-55 (1971).

[67] Loomes, Graham, and Ganna Progrebna, "Measuring Individual Risk Attitudes When Preferences Are Imprecise," *Economic Journal* 124: 569-593 (2014).

[68] Lucas, Robert E., Jr., "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy* 1: 19-46 (1976).

[69] Markowitz, Harry, "The Utility of Wealth," *Journal of Political Economy* 60: 151-158 (1952).

[70] Mosteller, Frederick, and Philip Nogee, "An Experimental Measurement of Utility," *Journal of Political Economy* 59: 371-404 (1951).

[71] Moyer, Robert S., and Thomas K. Landauer, "Time Required for Judgments of Numerical Inequality," *Nature* 215: 1519-1520 (1967).

[72] Muth, John F., "Rational Expectations and the Theory of Price Movements," *Econometrica* 29: 315-335 (1961).

[73] Myatt, David P., and Chris Wallace, "Endogenous Information Acquisition in Coordination Games," *Review of Economic Studies* 79: 340-374 (2012).

[74] Natenzon, Paulo, "Random Choice and Learning," *Journal of Political Economy* 127: 419-457 (2019).

[75] Netzer, Nick, Arthur Robson, Jakub Steiner, and Pavel Kocourek, "Risk Perception: Measurement and Aggregation," *J. European Economic Association* 23: 1309-1349 (2025).

[76] Nieder, Andreas, "The Neuronal Code for Number," *Nature Reviews Neuroscience* 17: 366-382 (2016).

[77] Odic, Darko, and Ariel Starr, "An Introduction to the Approximate Number System," *Child Development Perspectives* 12: 223-229 (2018).

[78] Oprea, Ryan, "Decisions Under Risk Are Decisions Under Complexity," *American Economic Review* 114: 3789-3811 (2024).

[79] Oprea, Ryan, and Ferdinand M. Vieider, "Minding the Gap: On the Origins of Probability Weighting and the Description-Experience Gap," working paper, UC Santa Barbara, February 2026.

[80] Petzschner, Frederike H., Stefan Glasauer, and Klaas E. Stephan, "A Bayesian Perspective on Magnitude Estimation," *Trends in Cognitive Sciences* 19: 285-293 (2015).

[81] Phillips, Lawrence D., and Ward Edwards, "Conservatism in a Simple Probability Inference Task," *Journal of Experimental Psychology* 72: 346-354 (1966).

[82] Piazza, Manuela, "Neurocognitive Start-Up Tools for Symbolic Number Representations," *Trends in Cogntive Sciences* 14: 542-551 (2010).

[83] Piazza, Manuela, Philippe Pinel, Denis Le Bihan, and Stanislaus Dehaene, "A Magnitude Code Common to Numerosities and Number Symbols in Human Intraparietal Sulcus," *Neuron* 53: 293-305 (2007).

[84] Plott, Charles R., "Rational Individual Behavior in Markets and Social Choice Processes: The Discovered Preference Hypothesis," in K.J. Arrow, E. Colombatto, M. Perleman, and C. Schmidt, eds., *Rational Foundations of Economic Behavior,* London: Macmillan, 1996.

[85] Porcelli, Anthony J., and Mauricio R. Delgado, "Acute Stress Modulates Risk Taking in Financial Decision Making," *Psychological Science* 20: 278-283 (2009).

[86] Prat-Carrabin, Arthur, and Michael Woodford, "Endogenous Precision of the Number Sense," *eLife* 13:RP101277 (2026).

[87] Prelec, Drazen, "The Probability Weighting Function," *Econometrica* 66: 497-527 (1998).

[88] Rubinstein, Ariel, "Response Time and Decision Making: An Experimental Study," *Judgment and Decision Making* 8(5): 540-551 (2013).

[89] Sargent, Thomas J., "Lucas' Critique," in *Macroeconomic Theory,* 2d ed., Orlando: Academic Press, 1987, pp. 397-398.

[90] Scholten, Mark, and Daniel Read, "Prospect Theory and the 'Forgotten' Fourfold Pattern of Risk Preferences," *Journal of Risk and Uncertainty* 48: 67-83 (2014).

[91] Sims, Christopher A., "Implications of Rational Inattention," *Journal of Monetary Economics* 50: 665-690 (2003).

[92] Steiner, Jakub, and Colin Stewart, "Perceiving Prospects Properly," *American Economic Review* 106: 1601-1631 (2016).

[93] Stott, Henry P., "Cumulative Prospect Theory's Functional Menagerie," *Journal of Risk and Uncertainty* 32: 101-130 (2006).

[94] Tong, Xin, Thi Thu Uyen Hoang, Xue-Xin Wei, and Michael Hahn, "The Bayesian Origin of the Probability Weighting Function in Human Representation of Probabilities," preprint posted on *arXiv,* January 4, 2026.

[95] Tversky, Amos, and Daniel Kahneman, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5: 297-323 (1992).

[96] Van de Kuilen, Gijs, "Subjective Probability Weighting and the Discovered Preference Hypothesis," *Theory and Decision* 67: 1-22 (2009).

[97] Van de Kuilen, Gijs, and Peter P. Wakker, "Learning in the Allais Paradox," *Journal of Risk and Uncertainty* 33: 155-164 (2006).

[98] Van den Berg, Ronald, and Wei Ji Ma, "A Resource-Rational Theory of Set-Size Effects in Human Visual Working Memory," *eLife* 7:e34963 (2018).

[99] Verguts, Tom, and Wim Fias, "Representation of Number in Animals and Humans: A Neural Model," *Journal of Cognitive Neuroscience* 16: 1493-1504 (2004).

[100] Vieider, Ferdinand M., "Decisions Under Uncertainty as Bayesian Inference on Choice Options," *Management Science* 70: 9014-9030 (2024).

[101] Vieider, Ferdinand M., "Cognitive Foundations of Delay-Discounting," working paper, RISL$\alpha\beta$, Ghent University, September 2025.

[102] Wakker, Peter P., "The History of Prospect Theory for Risk and Its Future for Ambiguity: Interactions Between Economists and Psychologists," working paper, Erasmus University Rotterdam, September 2025.

[103] Wei, Xue-Xin, and Alan A. Stocker, "A Bayesian Observer Model Constrained by Efficient Coding Can Explain 'Anti-Bayesian' Percepts," *Nature Neuroscience* 18: 1509-1517 (2015).

[104] Wei, Xue-Xin, and Alan A. Stocker, "Lawful Relation Between Perceptual Bias and Discriminability," *Proceedings of the National Academy of Sciences* 114: 10244-10249 (2017).

[105] Woodford, Michael, "Imperfect Common Knowledge and the Effects of Monetary Policy," in P. Aghion *et al, Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps,* Princeton University Press, 2003.

[106] Woodford, Michael, "Prospect Theory as Efficient Perceptual Distortion," *American Economic Review* 102(3): 1-8 (2012).

[107] Woodford, Michael, "Individualistic Welfare Analysis in the Age of Behavioral Science," *Capitalism and Society,* vol. 13, issue 1, article 3, posted online October 15, 2018.

[108] Woodford, Michael, "Modeling Imprecision in Perception, Valuation and Choice," *Annual Review of Economics* 12: 579-601 (2020).

[109] Xiang, Yang, Thomas Graeber, Benjamin Enke, and Samuel Gershman, "Bayesian Signatures of Confidence and Central Tendency in Perceptual Judgments," *Attention, Perception & Psychophysics* 83: 3024-3034 (2021).

[110] Young, Diana L., Adam S. Goodie, Daniel B. Hall, and Eric Wu, "Decision Making Under Time Pressure, Modeled in a Prospect Theory Framework," *Organizational Behavior and Human Decision Processes* 118: 179-188 (2012).

[111] Zhang, Hang, and Laurence T. Maloney, "Ubiquitous Log Odds: A Common Representation of Probability and Frequency Distortion in Perception, Action and Cognition," *Frontiers in Neuroscience* 6, article 1 (2012).

[112] Zhang, Hang, Xiangjuan Ren, and Laurence T. Maloney, "The Bounded Rationality of Probability Distortion," *Proceedings of the National Academy of Sciences* 117: 22024-22034 (2020).

[113] Zhou, Wenting, and John Hey, "Context Matters," *Experimental Economics* 21: 723-756 (2018).

ONLINE APPENDIX

Khaw, Li, and Woodford, "Cognitive Imprecision and Stake-Dependent Risk Attitudes"

A Details of the Experimental Design

Here we further discuss the exact protocol used in the experiment, and the incentives that it created, to the extent that the protocol was correctly understood by our subjects.

A.1 Bidding and Incentives

Each subject began with an endowment of US$30, that they could use to bid in any of the rounds. Any part of the endowment that was not spent, plus any gains or losses from the outcome of the lottery, would be taken home by the subject as payment for their participation in the experiment. While each subject bid on many successive lotteries (either 400 or 640, as explained further below), only one of these (randomly selected at the end of the experiment) would be used as the basis for their payment.

In the case of a "gain trial" (one with $X > 0$), the slider position indicated a bid $C \geq 0$ that the subject was willing to pay to receive the outcome of the lottery; the slider allowed arbitrary bids between zero and a maximum of $30 (i.e., all of their endowment). In the event that this trial was selected for payment, a computer bid B was generated as an independent draw from a uniform distribution on the interval [0, 30]. The "winner" of the BDM auction was then determined by which of the bids was higher. If the subject "won" (i.e., if $C > B$), then the subject would pay B (as in a second-price auction) and also receive the outcome of the lottery (an addition to their payment of $X > 0$ with probability p, but an addition of zero with probability $1 - p$). Hence they would either take home a payment of $30 - B$ or $30 - B + X$; in either case, the payment would be positive, since $B < 30$ with probability 1. If instead the subject "lost" (because $C < B$), they would simply take home their endowment of $30.

In the case of a "loss trial" (one with $X < 0$), instead, the slider position $|C|$ indicated the magnitude of the subject's negative bid ($C = -|C|$). A bid $|C|$ meant that the subject was willing to pay $|C|$ in order to avoid suffering the loss specified by the outcome of the lottery. In the event that this trial was selected for payment, a computer bid B was again generated as an independent draw from the uniform distribution on [0, 30], and the "winner" was again determined by whether B or $|C|$ was larger. If the subject "won" ($|C| > B$), then they would pay B, but avoid having to accept the outcome of the lottery; thus their payment in this case would be $30 - B$. If they "lost" ($|C| < B$), they would suffer the loss determined by the lottery, so that their payment would be $30 + X = 30 - |X|$ with probability p, and $30 with probability $1 - p$. Since $|X|$ was always less than $30 (never larger than $22.20), and B was less than $30 with probability 1, this would again necessarily result in a positive payment for participation in the experiment.

A.2 Losses from Inaccurate Bidding

We can now explain why minimization of the expected loss

$$\mathrm{E}[(C - pX)^2 \,|\boldsymbol{r}] \tag{A.1}$$

corresponds to maximization of the DM's expected financial wealth, given the financial incentives provided in our experiment. Consider first the case of a "gain trial" in which $X > 0$. Conditional on this trial being the one on which payment is based, the subject's expected[70] net addition to their payment (i.e., their expected net payment in excess of the \$30 endowment) from bidding an amount $0 \leq C \leq 30$ will equal

$$\frac{1}{30}\int_0^C (pX - B)\,dB \;=\; \frac{1}{30}\left[pX \cdot C \;-\; \frac{1}{2}C^2\right].$$

The maximum achievable value of this quantity (under full-information optimal bidding) is $(pX)^2/60$. The amount by which the maximum achievable expected payment exceeds the expected payment obtained by bidding C is thus equal to $L(C;\, pX)$, where we define

$$L(C;\, V) \;\equiv\; \frac{1}{60}(C \;-\; V)^2. \tag{A.2}$$

Now consider instead the case of a "loss trial" in which $X < 0$. Conditional on this trial being the one on which payment is based, the subject's expected net addition to their payment from a bid $0 \leq |C| \leq 30$ will be

$$\frac{1}{30}\int_0^{|C|} (-B)\,dB \;+\; \frac{1}{30}\int_{|C|}^{30} pX \cdot dB \;\;=\;\; \frac{1}{30}\left[pX \cdot (30 - |C|) \;-\; \frac{1}{2}|C|^2\right].$$

The maximum achievable value of this quantity is $pX + (pX)^2/60$. The amount by which the maximum achievable expected payment exceeds the expected payment obtained by bidding C is thus equal to

$$\frac{1}{30}\left[\frac{1}{2}(|C| \;+\; pX)^2\right].$$

But since we define this as a negative bid ($C = -|C|$), the loss from bidding C is again equal to $L(C;\, pX)$, using the definition (A.2). (In this case, both the bid C and the correct valuation V are negative.)

Thus an optimal bidding rule (neglecting any cognitive costs of producing the internal representation) is one that chooses a bid C (or a probability distribution for such bids), given the internal representation $\mathbf{r}$ of the decision situation, so as to minimize the expected loss

$$\mathrm{E}[L(C;\, pX)\,|\mathbf{r}]. \tag{A.3}$$

And minimization of (A.3) is equivalent to minimization of the objective (A.1) assumed in the main text. (Our omission of the prefactor 1/60, to simplify the notation, only affects the units in which the precision cost parameter A is expressed.)

[70] Here we mean the mathematical expectation of this quantity under our experimental protocol, which need not correspond to the subjective expectation of a subject. It is this mathematical expectation that we use to determine that a particular bidding rule would be optimal for a DM facing the decision problem posed in our experiment.

group	members	number of trials	values of p
1	1-6	400	0.1, 0.4, 0.6, 0.8, 0.9
2	13-15	400	0.1, 0.3, 0.5, 0.7, 0.9
3	16	400	0.05, 0.1, 0.5, 0.9, 0.95
4	17-19	400	0.05, 0.3, 0.5, 0.7, 0.95
5	7-12, 20-28	640	0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 0.9, 0.95

Table 3: Number of trials and values of p used for different groups of subjects.

A.3 Probabilities Used in the Lotteries

As explained in the main text, each subject was asked to evaluate a set of lotteries $(X; p)$, where both p and X are drawn from a finite set of possibilities. Each of the finite set of values for p (for that subject) was paired with each of the finite set of values for X, and each of the pairs $(X; p)$ that occurred for a given subject were presented equally often (8 times over the course of the session). The different lotteries $(X; p)$ were presented in a random order.

However, the finite set of values p that were used was different for different groups of subjects, as indicated in Table 3. The full set of 11 different probabilities were not used with any of the subjects; this allowed us to have multiple repetitions of the same problem for each of the subjects, in order to obtain a clear measure of trial-to-trial variability in the subject's response to each problem, without requiring excessively long experimental sessions.

The subjects are classified in the table as members of one or another of five groups, according to the set of lotteries presented to them. (One group, group 3, consists of only a single subject, subject 16.) In section 4.3 of the main text (and in more detail below, in Appendix section F), we classify subjects into two larger groups, the 400-trial subjects (the union of groups 1-4 in Table 3) and the 640-trial subjects (group 5). Note that while the 640-trial subjects all faced the same set of lotteries, the 400-trial subjects did not; each of these evaluated a set of lotteries using only five values of p, but the values of p used were different across the four groups of 400-trial subjects. We do not, however, estimate separate model parameters for the individual groups of 400-trial subjects, given that (at least in the case of groups 2, 3, and 4) there are only a few subjects in each group.

A.4 Zero Bids

When fitting our theoretical model to the experimental data, we exclude the bids which are equal to zero (the leftmost possible position of the slider), since, as explained in the main text, we regard this as declining to bid on that lottery. Here we provide additional information about the occurrence of these zero bids. Zero bids were more common among the subjects in the 640-trial group (who, as discussed further in Appendix section F, also displayed more signs of inattentiveness in other respects). The 12 non-excluded 640-trial subjects submitted zero bids on a total of 160 trials, or about 1.7 percent of all trials. Zero bids were instead relatively rare for the 400-trial subjects, who submitted only 15 such bids (less than 0.3 percent of their trials).

Zero bids also occurred much more frequently for some lotteries than for others, as shown by the "heat map" in Figure 10. Zero bids are most likely to occur when p or X (or both) are small. As the figure illustrates, most of the zero bids were submitted for lotteries with an EV of less than 3 dollars (in absolute value), meaning that the optimal bid would have been in the left-most 10 percent of the range of the slider. Many are in cases where the EV is not much more than a dollar (in absolute value). Zero bids were also somewhat more common in the case of lotteries involving losses: 60 percent of the zero bids occur in these cases, even though an equal number of lotteries involving losses and gains were presented to the subjects.[71] Zero bids were especially common in the case of lotteries involving losses and only a small probability ($p = 0.05$) of a non-zero loss; in this case, zero bids were submitted on 6.8 percent of all trials.

We assume that the decision whether to bother to submit a (non-zero) bid is based on a cursory inspection of the terms of the lottery $(X;\, p)$. This can be modeled as a decision rule conditioned on some noisy internal representation of the information $(X;\, p)$, though the information used for this first-stage decision need not be the same internal representations (r_p, r_x) that are used to choose a non-zero bid in the second stage (when it is reached). After all, we suppose that declining to submit a bid allows a saving of cognitive effort of some kind; this might mean not having to retrieve the noisy representations (r_p, r_x) that are instead needed if the DM chooses to submit a bid.[72]

Given the first-stage noisy internal representation and the first-stage decision rule, a DM has some probability $s(p, X)$ of choosing to submit a non-zero bid on a trial when the lottery is $(X;\, p)$.[73] The DM's prior in the second stage (when it is reached) should then depend on this selection effect. If $\pi(p, X)$ represents the distribution from which the experimenter draws values of $(X;\, p)$, then the DM's second-stage prior should be given by

$$\tilde{\pi}(p, X) \;=\; \frac{\pi(p, X)s(p, X)}{\mathrm{E}_{\pi}[s]}.$$

However, we simply take the second-stage prior $\tilde{\pi}(p, X)$ as given in our analysis of the second-stage problem. We estimate the parameters of the second-stage prior so as to fit as well as possible the empirically observed frequency distribution of lotteries $(X;\, p)$ that reach the second stage. Thus the observed pattern of selection of the lotteries for which the second stage is reached is taken into account, but we have no need (for our purposes here) to estimate a model of the first-stage decision. This is left for future study.

[71] This represents a departure from the symmetry of behavior in the gain and loss domains to which our data on non-zero bids by the non-excluded subjects largely conform. For example, we have shown in Table 1 that if the BIC is used as a basis for model comparison, the symmetric model is preferred to the unrestricted model, and the symmetric affine model is similarly preferred to the general affine model. But another suggestion that lotteries involving losses are more difficult to value, at least for some subjects, is the fact that three experimental subjects (subjects 9, 11 and 19) all seemed to have considerable difficulty understanding how to bid in the case of lotteries involving losses (their bids failed to increase monotonically with either p or $|X|$ to any appreciable degree), while only one of these (subject 11) had similar difficulty making minimally sensible bids in the case of lotteries involving gains. (The peculiarity of the bids of these subjects is discussed further in the September 2022 draft of NBER Working Paper no. 30417, Appendix, section D.2.)

[72] Similarly, we assume two distinct information structures (internal representations), each with a separate information cost, for the two stages of the decision problem in Khaw *et al.* (2017).

[73] An empirical measure of this probability is given by one minus the fraction shown in Figure 10 for each lottery $(X;\, p)$.

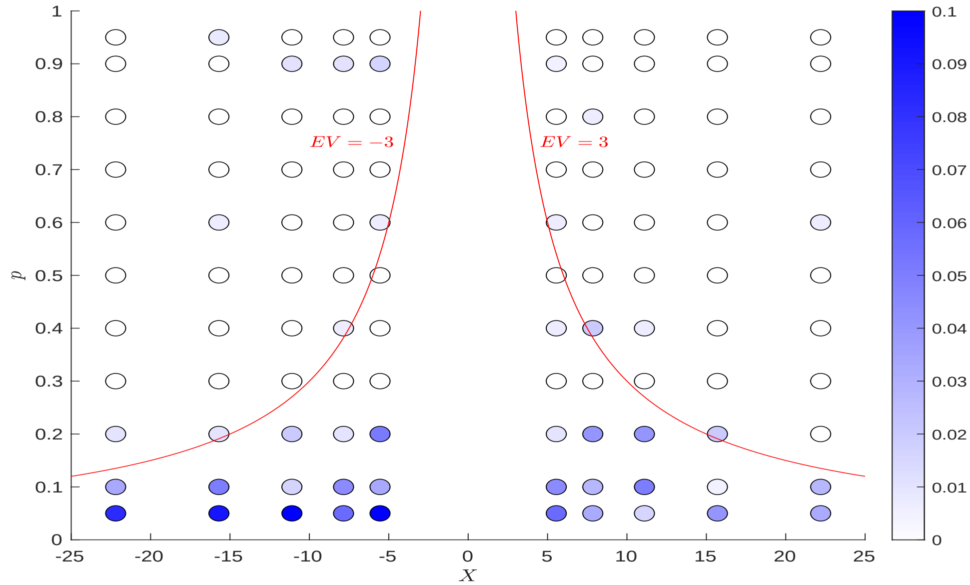


Figure 10: The fraction of zero bids for each of the lotteries $(X; p)$ that are presented to subjects. (Color code is explained by the scale at the right.)

B The Cognitive Cost of Precision in the Representation of Monetary Amounts

Our baseline model assumes that is possible to vary the precision with which monetary amounts are represented, subject to a cost that is linearly increasing in precision, as specified in (2.12). This specification of the cost function has a simple interpretation. Suppose that the magnitude $|X|$ is internally represented by a random quantity that evolves according to a Brownian motion, with a drift equal to $\log|X|$ and an instantaneous variance $\sigma^2 > 0$ that is independent of $|X|$. (It suffices for our argument that the drift be an affine function of $\log|X|$, but the calculations are simplified by assuming that the drift is simply equal to $\log|X|$. The assumption that $y_0 = 0$ is also purely to simplify the algebra.) This process y_t is allowed to evolve for some length of time $\tau > 0$, starting from an initial value $y_0 = 0$; the final value y_τ constitutes the internal representation. Diffusion processes of this kind are often used to model the randomness in sensory perception and memory retrieval.[74]

Equivalently, we may treat the value $r_x \equiv y_\tau/\tau$ as the internal representation, as this variable contains the same information as y_τ. Under this assumption, the internal representation has the distribution specified in (2.11), where $\nu_x^2 = \sigma^2/\tau$. Note that the precision of such a representation can be varied by varying τ, the length of time for which the process y_t is allowed to evolve. Moreover, successive increments of the Brownian motion are independent random variables (with a common distribution that depends on the magnitude $|X|$); these

[74]See Gold and Heekeren (2014) for a review. Heng *et al.* (2025) use a process of this kind to model the internal representation of positive numbers presented as arrays of dots, and show that the assumption of precision increasing linearly with time fits well the way that the distribution of errors in numerosity estimation varies with viewing time.

can be thought of as repeated noisy "readings" of the value of $|X|$.[75] If we suppose that each repeated "reading" has a separate (and identical) psychic cost, then the total cost should be linear in τ (and so proportional to the total number of independent "readings"). This implies a cost of precision of the form assumed in (2.12).

C The Log-Odds Model of Noisy Coding of Probabilities

Here we provide further justification for our interest in the model (3.1) for the noisy internal representation of the relative probability of the two possible outcomes of a lottery.

C.1 Consistency with the Logarithmic Model of Encoding of Positive Magnitudes

We first note that our model (3.1) of the imprecise representation of probability information is closely related to the way in which we model the imprecise representation of numbers, in our discussion of the internal representation of the monetary payoffs offered by a lottery.

Suppose that the relative probability of the two possible outcomes is displayed to a subject by the relative size of two magnitudes, X_1 and X_2, proportional to the probabilities of the two outcomes. (In the case of our experiment, X_1 and X_2 could be the lengths of the two bars corresponding to the probabilities of the two outcomes, as shown in Figure 2.) And suppose that each of these magnitudes is independently encoded by a noisy internal representation, where

$$r_j \;\sim\; N(\log X_j,\, \nu_p^2), \qquad j \;=\; 1, 2,$$

as specified for the monetary amounts in (2.11). (Note that this would also be a common model of imprecision in visual perception of length.) Finally, suppose that judgments about the relative probability of the two outcomes are based purely on the *difference* between these two internal representations, $r_p \equiv r_1 - r_2$. In this case, the conditional distribution of the internal representation r_p of the relative odds will be of the form (3.1), where p in this expression means the probability of outcome 1, and $\nu_z^2 = 2\nu_p^2$. Note, however, that our conclusions in our baseline of model of lottery valuation depend only on assuming (3.1), and not on this particular interpretation of how the internal representation of the relative odds may be constructed.

Another reason for proposing (3.1) as a model of imprecision in the internal representation of probability information is the usefulness of a model of this kind in accounting for measured imprecision of people's judgments about probabilities, relative frequencies, and proportions, when these are presented visually or through a sample of instances (rather than with number symbols as in our experiment). The idea is parallel to our hypothesis about the encoding of numerical magnitudes: that the imprecision in the internal representation of numbers is the same when numbers are presented symbolically (as in our experiment) as in the better-studied case of judgments about numbers presented visually (numbers represented by the length of a bar, or the number of items in an array).

[75] Gold and Heekeren (2014) discuss the neural mechanisms that could implement such a process.

For example, Eckert *et al.* (2018) present evidence for a similar model of the discriminability of different relative frequencies of occurrence of two possible outcomes. They experimentally test the ability of both humans and chimpanzees to distinguish between two urns, one with a ratio $a_1 : a_2$ of outcomes of type 1 rather than of type 2, and another with a ratio $b_1 : b_2$ of the two possible outcomes, and find that the probability of correctly recognizing which offers the higher probability of outcome 1 is an increasing function of the "ratio of ratios" $(a_1/a_2)/(b_1/b_2)$. Equivalently, it is an increasing function of the difference in the log odds associated with the two urns,

$$\log \frac{a_1}{a_2} - \log \frac{b_1}{b_2}.$$

This is exactly the prediction of our model, if each relative frequency $p_i : (1 - p_i)$ is encoded by a noisy internal representation r_{pi} drawn from the distribution (3.1), and the DM's judgment is based on the relative size of r_{p1} and r_{p2}. Note also that Eckert *et al.* conclude from their findings "that intuitive statistical reasoning relies on the same cognitive mechanism that is used for comparing absolute quantities, namely the analogue magnitude system." Here by "the analogue magnitude system" they mean a system of imprecise semantic representation of natural numbers, with the property that "discriminability of two sets varies as a function of the ratio of the set sizes to be compared, independent of their absolute numerosity,"[76] as would be implied if the two numerical magnitudes are encoded logarithmically as specified in (2.11).

C.2 An Optimizing Model of Bias in the Estimation of Probabilities or Relative Frequencies

Additional (though slightly less direct) evidence in favor of a model of noisy internal representation of probabilities like (3.1) comes from studies in which subjects must produce an *estimate* of some probability or proportion, rather judging which of two probabilities is greater. While studies of biases in estimation (as opposed to discriminability) provide less direct evidence about the precision of the internal representations on which the estimates are based, they are arguably of more direct relevance to the cognitive task in our experiment (i.e., producing an estimate of the value of a lottery).

Studies of bias in the estimation of probabilities, relative frequencies, and proportions find that people's estimates are typically most accurate for probabilities near 0 or 1, but much less accurate for intermediate probabilities (Zhang and Maloney, 2012). Moreover, at least in cases where there are only two possible outcomes, and the distribution of values for the probability of the first outcome is symmetric around 0.5, the degree of estimation error is typically found to be symmetric around 0.5, as the specification (3.1) together with a hypothesis of Bayesian decoding would imply. In fact, Zhang and Maloney (2012) review a wide range of previous experiments requiring subjects to judge the relative frequency with which two outcomes occur — either when presented simultaneously (say, a visual image containing many randomly arranged dots of two different colors) or in sequence (say, a succession of letters that are either of one type or the other). They show that

[76] See Eckert *et al.* (2018, abstract and p. 100).

characteristically, the median estimate $\bar{p}$ is a function of the true probability (or relative frequency) p of the form

$$\log \frac{\bar{p}}{1-\bar{p}} \;=\; \gamma \, \log \frac{p}{1-p} \;+\; (1-\gamma) \, \log \frac{p_0}{1-p_0}, \tag{C.1}$$

for some "anchor" or reference probability p_0 and an adjustment coefficient that in most cases satisfies $0 < \gamma < 1$. The reference probability p_0 is different in different experiments, but Zhang and Maloney note that it is typically close to the average of the true values p used in the experimental trials. Here we show how such a pattern of bias can result from optimal Bayesian decoding of a noisy internal representation of the kind specified by (3.1).[77]

Bayesian decoding of the noisy internal representation can only be defined relative to a prior distribution of true values of p for which the subject's decision rule has been optimized. A hypothesis that is convenient for such calculations (and that delivers a linear-in-log-odds relationship, at least approximately) is to assume a logit-normal prior,

$$z \;\sim\; N(\mu_z, \sigma_z^2), \tag{C.2}$$

where we introduce the notation $z \equiv \log(p/1-p)$ for the log odds. In the case of such a prior, the posterior distribution for the log odds, conditional on the representation r_p, will be a Gaussian distribution

$$z \;\sim\; N(\hat{\mu}_z(r_p), \hat{\sigma}_z^2), \tag{C.3}$$

where

$$\hat{\mu}_z(r_p) \;=\; \mu_z + \left(\frac{\sigma_z^2}{\sigma_z^2 + \nu_z^2}\right)(r_p - \mu_z), \qquad \hat{\sigma}_z^{-2} \;=\; \sigma_z^{-2} + \nu_z^{-2}. \tag{C.4}$$

It is not entirely clear what objective should be maximized by subjects' responses in the experiments reviewed by Zhang and Maloney (2012), since the experiments are typically not incentivized (and of course, one might assume in any event that there should be important "psychic" benefits from accuracy in addition to any monetary rewards). One simple hypothesis might be that the subject's estimate $\hat{p}$ is the one implied by the maximum a posteriori (MAP) estimate of the log odds of the event conditional on an internal representation r_p with statistics of the kind proposed above.[78] In this case, the model predicts an estimate

$$\hat{p}(r_p) \;=\; \frac{e^{\hat{z}(r_p)}}{1 + e^{\hat{z}(r_p)}}, \tag{C.5}$$

where the estimated log odds are given by $\hat{z}(r_p) = \hat{\mu}_z(r_p)$, the function defined in (C.4). We obtain the same prediction if instead we suppose that a subject computes an estimate of

[77] Zhang *et al.* (2020) propose a related model, and fit it to a variety of experimental datasets, though their model of the noisy coding of probability information is more complex, and their model of estimation on the basis of the noisy internal representation is not fully Bayesian.

[78] This hypothesis is discussed mainly because it allows us a simple closed-form solution. However, in at least some experimental studies of probability estimation, subjects report their probability estimate in terms of log odds; see Phillips and Edwards (1966). And Zhang and Maloney (2012) argue that there is reason to believe that the brain represents probabilities in terms of log odds, so that probability estimates can be understood as resulting from intuitive calculations in terms of log odds.

the log odds given by the posterior mean value of z, and then converts this into an implied estimate for p using (C.5).

Then since $\hat{\mu}_z(r_p)$ is a monotonic function, and the estimate for p specified in (C.5) is also a monotonic function of the estimate for z, the median estimate of p is predicted to be

$$\bar{p} \;=\; \hat{p}(z) \;=\; \frac{e^{\hat{\mu}_z(z)}}{1+e^{\hat{\mu}_z(z)}}.$$

This implies that

$$\log\frac{\bar{p}}{1-\bar{p}} \;=\; \hat{\mu}_z(z),$$

which is a relation of the form (C.1), in which

$$\gamma \;=\; \hat{\gamma} \;\equiv\; \frac{\sigma_z^2}{\sigma_z^2+\nu_z^2}, \qquad \log\frac{p_0}{1-p_0} \;=\; \mu_z. \tag{C.6}$$

The average estimated log odds would thus be an increasing function of the true log odds, with a slope less than one, implying a conservative bias. Moreover, the cross-over value is predicted to be the probability corresponding to log odds of $z = \mu_z$: the mean of the possible log odds under the prior. Hence this kind of Bayesian model provides a potential explanation for the results summarized in Zhang and Maloney (2012).

An alternative behavioral model would assume that subjects' estimates of p correspond to the posterior mean value of p (rather than the value of p implied by the posterior mean value of z); that is, that $\hat{p} \;=\; \mathrm{E}[p\,|r_p]$. In this case, we cannot give an explicit analytical solution for $\hat{p}(r_p)$, but Daunizeau (2017) offers a "semi-analytical" solution which he shows numerically is quite accurate over a wide range of parameter values. Using this result, the posterior expected value $\hat{p}$ can be approximated by the value such that

$$\log\frac{\hat{p}}{1-\hat{p}} \;=\; \alpha\hat{\mu}_z(r_p), \tag{C.7}$$

where

$$\alpha \;=\; [1\,+\,a\hat{\sigma}_z^2]^{-1/2} \;<\; 1$$

and a is a constant equal to about 0.368. The median estimate of p should then satisfy

$$\log\frac{\bar{p}}{1-\bar{p}} \;=\; \alpha\hat{\mu}_z(z), \tag{C.8}$$

which is again a relation of the form (C.1), but now with

$$\gamma \;=\; \alpha\hat{\gamma}, \qquad \log\frac{p_0}{1-p_0} \;=\; \left(\frac{\alpha(1-\hat{\gamma})}{1-\alpha\hat{\gamma}}\right)\mu_z,$$

where $\hat{\gamma}$ is again defined as in (C.6).

Again we find (to an excellent degree of approximation) that the relationship between p and the median estimate $\bar{p}$ should be of the linear-in-log-odds form assumed in the regressions of Zhang and Maloney (2012). Again the average estimated log odds would thus be an increasing function of the true log odds, with a slope less than one, implying a conservative bias; and again the value of the log odds at which the cross-over from over-estimation to under-estimation should occur is an increasing function of μ_z (though deviating from even odds slightly less than does μ_z). The consistency of these results with the empirical evidence in Zhang and Maloney (2012) suggests that the model (3.1) of imprecise encoding of probability information is a realistic one.

Note that in an experiment like that of Enke and Graeber (2023), in which the magnitude $|X|$ is the same on all trials (with only p and the sign of X differing across trials), a model of bias in the estimation of probabilities directly implies a model of bias in lottery valuations. If $|X|$ is the same on all trials, and we assume a decision rule that is optimized for the distribution of lotteries actually encountered in the experiment, then there can be no posterior uncertainty about the value of $|X|$. Then if we ignore the issue of response error (analyzed in section D.2), the bidding rule that would maximize the DM's expected financial wealth will simply be

$$C \;=\; \mathrm{E}[p\,|r_p] \cdot X,$$

so that our model predicts

$$\log \frac{WTP}{EV} \;=\; \log \mathrm{E}[p\,|r_p] \;-\; \log p. \tag{C.9}$$

Thus the relative risk premium implied by subjects' bids should (according to our model) be purely a function of the bias in the optimal Bayesian estimate of p conditional on the noisy internal representation r_p of the relative probabilities.

In the case of a symmetric prior distribution (one in which the relative probabilities $(1-p, p)$ are exactly as likely as $(p, 1-p)$ for all p), we should have $\mu_z = 0$. Our results above then imply that p_0 should equal 0.5, and that we should observe that subjects' median bids should satisfy $|WTP| > |EV|$ for lotteries with $p < 0.5$ and $|WTP| < |EV|$ for lotteries with $p > 0.5$, in either the gain or loss domain, as Enke and Graeber (2023) find.[79]

Moreover, fixing the prior distribution of the probabilities, the size of these biases (i.e., the systematic departures from risk-neutral bidding) should depend only on σ_z^2, the degree of imprecision in the internal representation of probabilities. A larger value of σ_z^2 should increase $\hat{\sigma}_z^2$, and as a consequence should lower the value of α. It should also make $\hat{\mu}_z(z)$ closer to zero, for any value of z. Hence for both reasons, the median value $\bar{p}$ of the posterior mean estimate of p given by (C.8) should be closer to 0.5, for any true p, the larger is σ_z^2. This in turn means that for any $p \neq 0.5$, the size of the departure from risk-neutral bidding implied by (C.9) should be an increasing function of σ_z^2. This prediction is consistent both with the results of Enke and Graeber that show that subjects with higher reported cognitive uncertainty exhibit larger departures from risk-neutrality (in all four quadrants of the Tversky-Kahneman "fourfold pattern"), and with their demonstration that interventions that ought to reduce the precision of subjects' awareness of the value of p cause them to exhibit larger departures from risk-neutrality (again in all four quadrants).

We show how these predictions can be extended to the more general case in which there is cognitive uncertainty about the magnitude $|X|$ of the monetary payoff as well, and also

[79] The findings of Enke and Graeber, of course, are equally consistent with a model in which it is EV (rather than p) that is represented (or retrieved) with noise. But as discussed in section 4.2 of the main text, that alternative type of model of optimal adaptation to cognitive noise is much less successful at explaining the pattern of apparent risk attitudes in our experimental data.

derive the consequences of taking into account unavoidable response error, in the section that follows.

D Noisy Coding and Lottery Valuation: Derivations

Here we explain the details of the derivation of the theoretical model sketched in the main text. We begin with a complete derivation of the quantitative predictions of our baseline model, and then briefly discuss the predictions of the alternative models of noisy coding that are compared in Table 2.

D.1 Optimal Decoding of the Imprecise Representation of the Monetary Payoff

The optimal bid based on a given noisy representation of the decision problem depends on the posterior distribution over possible decision problems implied by that representation. We begin by discussing optimal (Bayesian) inference about the magnitude $|X|$ of the monetary payoff promised by the lottery. The quantities $(X; p)$ that specify the decision problem on a given trial have noisy internal representations (r_p, r_x), the conditional distributions of which are given by

$$r_p|p \;\sim\; N(\log(p/1-p),\, \nu_z^2), \qquad r_x|(r_p, X) \;\sim\; N(\log X,\, \nu_x^2(r_p)),$$

where the function $\nu_x^2(r_p)$ is to be optimized (but is taken as given in this section). Note that the conditional distribution of r_p is independent of the size of $|X|$, and that the conditional distribution of r_x depends on the value of p only through its internal representation r_p. We can view r_p as being determined first, in a way that depends only on the value of p; the internal representation r_x is then determined by $|X|$, but in a way that can depend on the already encoded value r_p.

The DM is assumed to correctly recognize the sign of X on a given trial, but will have a non-degenerate posterior distribution over possible exact values of the lottery characteristics. The Bayesian posterior conditional on a particular imprecise internal representation (r_p, r_x) depends on the prior distribution from which the true values $(X; p)$ are expected to have been drawn. We suppose that p and $|X|$ are independent random variables, with a prior distribution for $|X|$ given by

$$\log|X| \;\sim\; N(\mu_x,\, \sigma_x^2).$$

(The conclusions in this section are independent of what we assume about the prior distribution for p, other than that the two variables are distributed independently of one another.)

Under the assumption of a log-normal prior for $|X|$, the posterior for $|X|$ is also log-normal. It follows that

$$\mathrm{E}[|X|\,|r_p, r_x] \;=\; \exp[(1-\gamma_x(r_p))\mu_x \;+\; \gamma_x(r_p)r_x \;+\; \frac{1}{2}(1-\gamma_x(r_p))\sigma_x^2], \tag{D.1}$$

where

$$\gamma_x(r_p) \;\equiv\; \frac{\sigma_x^2}{\sigma_x^2 + \nu_x^2(r_p)}$$

is a quantity satisfying $0 < \gamma_x(r_p) < 1$ that can be different for each r_p. It similarly follows that

$$\mathrm{E}[X^2 \,|r_p, r_x] \;=\; \exp[2(1-\gamma_x(r_p))\mu_x \;+\; 2\gamma_x(r_p)r_x \;+\; 2(1-\gamma_x(r_p))\sigma_x^2].$$

D.2 Implications of Cognitive Noise for Optimal Bidding

If C could be chosen with precision, given an internal representation (r_p, r_x), the solution to this problem would be to choose the bid specified in (2.6). That is, the optimal bid would simply be the mean of the Bayesian posterior distribution for the true EV of the lottery, conditional on the imprecise internal representation of the problem. It follows from our results above that in the case of a "gain trial" with $X > 0$, the optimal bid will be a quantity $C > 0$ such that

$$\begin{aligned}\log C &= \log \mathrm{E}[p\,|r_p] \;+\; \log \mathrm{E}[X\,|r_p, r_x] \\ &= \log \mathrm{E}[p\,|r_p] \;+\; (1-\gamma_x(r_p))\mu_x \;+\; \gamma_x(r_p)r_x \;+\; \frac{1}{2}(1-\gamma_x(r_p))\sigma_x^2. \qquad \text{(D.2)}\end{aligned}$$

However, because of the presence of unavoidable response error, it is only the mean of the distribution (3.5) that can be chosen as a function of $\mathbf{r}$, and not the value of C that will be bid on any given trial. If response error were assumed to be additive, i.e., if instead of (3.5) we were to assume that

$$C \;\sim\; N(f(r_p, r_x),\, \nu_c^2),$$

a "certainty equivalence" result would obtain: the right-hand side of (D.2) would still be the log of the optimal choice for the "target" (or intended bid) f, though the actual bid would equal this plus a mean-zero noise term. But because we have (more accurately, in our view) specified a multiplicative response error in (3.5), the optimal solution is more complex.

As explained in the main text, our model of imprecise response selection implies that the DM's (unsigned) bid $|C|$ will be given by

$$\log |C| \;=\; f(r_p, r_x) \;-\; \epsilon_c, \qquad \text{(D.3)}$$

where

$$\epsilon_c \;\sim\; N(0,\, \nu_c^2)$$

is distributed independently of r_p and r_x. We now consider the optimal choice of the "target" function f.

For each possible internal representation (r_p, r_x), we can write a separate optimization problem: choose $f(r_p, r_x)$ to minimize the expected loss

$$\begin{aligned}\mathrm{E}[(C - pX)^2\,|r_p, r_x] &= \mathrm{E}[C^2\,|r_p, r_x] \;-\; 2\mathrm{E}[CpX\,|r_p, r_x] \;+\; \mathrm{E}[p^2X^2\,|r_p, r_x] \\ &= \mathrm{E}[\exp(2\epsilon_c)] \cdot \exp(2f(r_p, r_x)) \\ &\quad - 2\mathrm{E}[\exp(\epsilon_c)] \cdot \exp(f(r_p, r_x)) \cdot \mathrm{E}[p\,|r_p] \cdot \mathrm{E}[X\,|r_p, r_x] \\ &\quad + \mathrm{E}[p^2\,|r_p] \cdot \mathrm{E}[X^2\,|r_p, r_x],\end{aligned}$$

where we have used (D.3) to substitute for C as a function of r_p, r_x, and ϵ_c.

This is a quadratic function of $\exp(f(r_p, r_x))$. Moreover, since

$$\mathrm{E}[\exp(2\epsilon_c)] \;=\; \exp(2\nu_c^2) \;>\; 0,$$

the expected loss is a strictly convex function, with a unique minimum when

$$\mathrm{E}[\exp(2\epsilon_c)]\exp(f(r_p, r_x)) \;=\; \mathrm{E}[\exp(\epsilon_c)] \cdot \mathrm{E}[p\,|r_p] \cdot \mathrm{E}[X\,|r_p, r_x].$$

Using the fact that both X and ϵ_c are log-normally distributed (conditional on r_p, r_x), we can express the optimal choice of f as

$$f(r_p, r_x) \;=\; \log \mathrm{E}[p\,|r_p] \;+\; (1-\gamma_x(r_p))[\mu_x + \frac{1}{2}\sigma_x^2] \;+\; \gamma_x(r_p) r_x \;-\; \frac{3}{2}\nu_c^2.$$

When f is chosen in this way, the minimized value of the expected loss is

$$\begin{aligned}\mathrm{E}[(C-pX)^2\,|r_p, r_x] \;\;&=\;\; \exp(2(1-\gamma_x(r_p))[\mu_x + \frac{1}{2}\sigma_x^2] \;+\; 2\gamma_x(r_p) r_x) \cdot \\ &\Big\{\exp((1-\gamma_x(r_p))\sigma_x^2)\mathrm{E}[p^2\,|r_p] \;-\; \exp(-\nu_c^2)\mathrm{E}[p\,|r_p]^2\Big\}. \qquad \text{(D.4)}\end{aligned}$$

Substitution of this solution into (D.3) implies that the equation

$$\begin{aligned}\log C \;-\; \log(pX) \;\;&=\;\; \big(\log \mathrm{E}[p\,|r_p] - \log p\big) \;+\; (1-\gamma_x(r_p))[\mu_x + \frac{1}{2}\sigma_x^2 - \log X] \\ &+ \gamma_x(r_p)[r_x - \log X] \;-\; \frac{3}{2}\nu_c^2 \;+\; \epsilon_c\end{aligned}$$

gives the predicted value of $\log(WTP/EV)$ in the case of any given lottery $(X;\,p)$, any given internal representation (r_p, r_x), and any given realization of the response noise ϵ_c. Integrating over the conditional distributions of the random variables (r_p, r_x, ϵ_c) in the case of a given lottery $(X;\,p)$, we obtain the prediction that

$$\mathrm{E}[\log(C/pX)\,|p, X] \;=\; \alpha_p \;+\; \beta_p \log X, \qquad \text{(D.5)}$$

where the coefficients

$$\alpha_p \;\equiv\; \mathrm{E}\big[\log \mathrm{E}[p\,|r_p] - \log p\;|p\big] \;+\; (1-\gamma_p)[\mu_x + \frac{1}{2}\sigma_x^2] \;-\; \frac{3}{2}\nu_c^2, \qquad \text{(D.6)}$$

$$\beta_p \;\equiv\; -(1-\gamma_p),$$

$$\gamma_p \;\equiv\; \mathrm{E}[\gamma_x(r_p)\,|p]$$

all depend on the value of p but are independent of X.

Since $0 < \gamma_x(r_p) < 1$ for each possible value of r_p, it follows that $0 < \gamma_p < 1$ for each value of p, and hence that $-1 < \beta_p < 0$ for each p. We thus conclude that for any lottery $(X;\,p)$, the predicted distribution of values for WTP (i.e., the distribution of the random variable C in (D.5)) is such that the mean value of $\log(WTP/EV)$ should be an affine function of $\log X$, with a slope and intercept that can vary with p. Furthermore, for each value of p, the slope must satisfy $-1 < \beta_p < 0$. These predictions are tested in the way discussed in section 1.3 of the main text.

In the case that X is negative (the lottery offers a random loss rather than a random gain), we suppose that p and the magnitude $|X|$ are encoded with noise in the same way (and with the same parameters) as is specified above in the case that X is positive. The optimal bid in this case will obviously be negative; we assume that in the case of a negative bid C, the absolute value $|C|$ will again be given by the right-hand side of (D.3), just as in the case of a positive bid. The optimal function $f(r_p, r_x)$ will then be exactly the same as in the derivation above. We conclude that the distribution of values for C/pX will be exactly the same function of p and $|X|$ as in the case where X is positive. In particular, (D.5) will again hold, except with $\log X$ replaced by $\log |X|$ on the right-hand side; the coefficients α_p, β_p will be the same functions of p as in the case of random gains. This prediction is also tested in the way discussed in the main text.

D.3 Endogenous Encoding Precision

We turn now to the way in which the coefficients α_p, β_p are predicted to vary with p. This depends on what we assume about the noisy encoding of p, and about the prior over values of p for which the decision rule is optimized; but it also depends on what we assume about how $\nu_x^2(r_p)$ varies with r_p. We suppose that the latter function is endogenously determined, so as to maximize the accuracy of bidding subject to a cost of encoding precision, as discussed in the main text.

Note that our model of noisy coding implies that conditional on the value of r_p, the distribution of r_x is

$$r_x | r_p \;\sim\; N(\mu_x,\, \sigma_x^2 + \nu_x^2(r_p)), \tag{D.7}$$

from which it follows that

$$2\gamma_x(r_p) r_x | r_p \;\sim\; N(2\gamma_x(r_p)\mu_x,\, 4\gamma_x(r_p)\sigma_x^2).$$

Thus exponentiation of this variable results in a log-normal random variable, with mean

$$\mathrm{E}[\exp(2\gamma_x(r_p) r_x)\,|r_p] \;=\; \exp(2\gamma_x(r_p)\mu_x \;+\; 2\gamma_x(r_p)\sigma_x^2).$$

Using this result, we can then integrate (D.4) over the distribution (D.7) for r_x to obtain

$$\mathrm{E}[(C - pX)^2\,|r_p] \;=\; \exp(2(\mu_x + \frac{1}{2}\sigma_x^2)) \cdot \left\{\exp(\sigma_x^2)\mathrm{E}[p^2\,|r_p] \;-\; \exp(\gamma_x(r_p)\sigma_x^2 - \nu_c^2)\mathrm{E}[p\,|r_p]^2\right\}.$$

Thus we can write

$$\mathrm{E}[(C - pX)^2\,|r_p] \;=\; Z(r_p) \;-\; \Gamma\varphi(r_p) \cdot \exp(\gamma_x(r_p)\sigma_x^2), \tag{D.8}$$

where

$$\Gamma \;\equiv\; \exp(2(\mu_x + \frac{1}{2}\sigma_x^2) \;-\nu_c^2) \;>\; 0, \qquad \varphi(r_p) \;\equiv\; \mathrm{E}[p\,|r_p]^2 \;>\; 0,$$

and $Z(r_p)$ is a term that is independent of the choice of $\nu_x^2(r_p)$. We thus observe that for any r_p, the expected loss conditional on r_p is a monotonically decreasing function of $\gamma_x(r_p)$, and hence a monotonically increasing function of $\nu_x^2(r_p)$.

If the cost of greater precision in the encoding of X, in the same units as those in which $L(C; pX)$ is measured, is given by

$$\frac{A}{\nu_x^2} \;=\; \frac{A}{\sigma_x^2}\Big(\frac{\gamma_x}{1-\gamma_x}\Big),$$

as assumed in (2.12), then minimization of total costs (counting the cost of precision) requires that for each r_p, the value of $\gamma_x(r_p)$ be the solution to the problem

$$\min_{\gamma_x} F(\gamma_x;\, r_p) \;\equiv\; \frac{A}{\sigma_x^2}\Big(\frac{\gamma_x}{1-\gamma_x}\Big) \;-\; \Gamma\varphi(r_p)\cdot\exp(\gamma_x\sigma_x^2). \tag{D.9}$$

We further observe that

$$\frac{\partial F}{\partial \gamma_x} \;=\; \frac{A}{\sigma_x^2}\frac{1}{(1-\gamma_x)^2} \;-\; \Gamma\varphi(r_p)\sigma_x^2\cdot\exp(\gamma_x\sigma_x^2),$$

an expression that has a positive sign if and only if

$$\frac{\tilde{A}}{(1-\gamma_x)^2} \;>\; \varphi(r_p)\exp(\gamma_x\sigma_x^2), \tag{D.10}$$

where we now use the notation

$$\tilde{A} \;\equiv\; \frac{A}{\Gamma\sigma_x^4} \;>\; 0 \tag{D.11}$$

as an alternative parameterization of the size of the cost of precision. Taking the logarithm of both sides of the inequality (D.10), we see that

$$\frac{\partial F}{\partial \gamma_x} \;>\; 0 \;\Leftrightarrow\; G(\gamma_x;\, r_p) \;>\; 0,$$

where we define

$$G(\gamma_x;\, r_p) \;\equiv\; \log\tilde{A} \;-\; \log\varphi(r_p) \;-\; 2\log(1-\gamma_x) \;-\; \gamma_x\sigma_x^2. \tag{D.12}$$

We see from this that $F(\gamma_x;\, r_p)$ is a decreasing function of γ_x at $\gamma_x = 0$ if and only if

$$\tilde{A} \;<\; \varphi(r_p), \tag{D.13}$$

so that $G(0;\, r_p) \;<\; 0$. We also note that $F(\gamma_x;\, r_p)$ is an increasing function of γ_x as $\gamma \to 1$ (indeed, increasing without bound). Hence (D.13) is a sufficient condition for the existence of an interior solution to the problem (D.9) at some $0 < \gamma_x < 1$. Moreover, the function defined in (D.12) is a strictly convex function of γ_x; hence its graph can cross the line $G = 0$ for at most two values of γ_x, and then only if $G > 0$ at both extremes.

Thus if (D.13) holds, there must be exactly one solution to the first-order condition

$$G(\gamma_x;\, r_p) \;=\; 0. \tag{D.14}$$

In addition, we must have $G < 0$ for all smaller values of γ_x, while $G > 0$ for all greater values of γ_x. From this it follows that the solution to the FOC (D.14) must be the global minimum of the function F, and hence the solution to problem (D.9).

We also observe that the value of r_p affects this solution only through its effect on the value of $\varphi(r_p)$; thus we can solve for the optimal γ_x as a function of the value of $\varphi(r_p)$. When $\varphi(r_p)$ satisfies (D.13), so that we have an interior solution to the FOC, we can compute the derivative of γ_x with respect to changes in the value of $\varphi(r_p)$ through total differentiation of the FOC. It follows from (D.12) that

$$\frac{\partial G}{\partial \varphi} = -\frac{1}{\varphi} < 0, \qquad \frac{\partial G}{\partial \gamma_x} = \frac{2}{1-\gamma_x} - \sigma_x^2 > 0,$$

as long as $\sigma_x^2 \leq 2$. (This is in fact the case of interest for us, since in our experiment, the variance of $\log |X|$ is approximately 0.26.)

Then total differentiation of the FOC (D.14) implies that

$$\frac{d\gamma_x}{d\varphi(r_p)} = -\frac{\partial G/\partial \varphi}{\partial G/\partial \gamma_x} > 0.$$

It follows that the optimal solution for γ_x will be a monotonically increasing function of $\varphi(r_p)$, with $\gamma_x \to 0$ as $\varphi \to \tilde{A}$ and $\gamma_x \to 1$ as $\varphi \to \infty$. Or equivalently, the optimal solution for ν_x^2 will be a monotonically decreasing function of $\varphi(r_p)$, with $\nu_x^2 \to \infty$ as $\varphi \to \tilde{A}$ and $\nu_x^2 \to 0$ as $\varphi \to \infty$.

Let us now consider the alternative case in which $\varphi(r_p) \leq \tilde{A}$. In this case $G \geq 0$ when $\gamma_x = 0$, and since $\partial G/\partial \gamma_x > 0$ (again assuming that $\sigma_x^2 \leq 2$), it follows that $G > 0$ for all $\gamma_x > 0$. This implies that $\partial F/\partial \gamma_x > 0$ for all $\gamma_x > 0$, so that the solution to the problem (D.9) must be $\gamma_x = 0$ in all such cases. Thus we obtain a unique optimal solution for γ_x (and hence for ν_x^2) for any value of $\varphi(r_p)$. The optimal γ_x is a non-decreasing function of $\varphi(r_p)$: constant (and equal to zero) for all $0 \leq \varphi(r_p) \leq \tilde{A}$, and increasing for all $\varphi(r_p) > \tilde{A}$.

D.4 Alternative Models of Noisy Coding

In section 4.2 of the main text, we consider a variety of alternatives to the baseline cognitive noise model analyzed above. Here we briefly discuss how the quantitative predictions of each of these models are derived.

Exogenous precision. This model assumes as a constraint that $\nu_x(r_p) = \nu_x$ for all r_p, where ν_x is now a constant to be estimated. In this case, the expressions derived above continue to apply, but γ_x is now a quantity (between 0 and 1) independent of r_p. As a result, γ_p takes the same value (equal to γ_x) for all p, and its value can be calculated from the values of parameters σ_x and ν_x, independently of assumptions about the encoding and decoding of the probabilities. An important implication is that $\beta_p = -(1-\gamma_p)$ will be the same negative quantity for all values of p; thus stake-size effects are predicted to be the same for all values of p (contrary to what we, and others, find empirically).

Noiseless retrieval of monetary payoffs. This model assumes that $\nu_x = 0$; it is thus a special case of the exogenous noise model, in which ν_x is no longer a free parameter. In this special case, we have the stronger result that $\gamma_x = 1$ (again regardless of the value of r_p), and hence that $\beta_p = 0$ for all p. Thus this model implies that there should be no stake-size effects.

Noiseless retrieval of probabilities. This model assumes that $\nu_z = 0$, so that r_p can be identified with the value of p itself. In this case, condition (D.14) reduces to the equation

$$\frac{\tilde{A}}{(1-\gamma_p)^2} \;=\; p^2 \cdot \exp(\gamma_p \sigma_x^2).$$

This is an equation that implicitly defines the value of γ_p (and hence the value of β_p) for any value of p; it is no longer necessary to solve for $\gamma_x(r_p)$ for each of the possible r_p associated with a given probability p, and then numerically integrate over the distribution of such solutions in order to obtain a prediction for γ_p. Condition (D.6) yields an expression for α_p that can be solved in closed form, and that is the same for all p. And finally, we obtain a closed-form solution for the variance of the distribution of $\log(WTP)$ as well, namely

$$\mathrm{var}(\log(WTP)\,|p, X) \;=\; \gamma_p^2 \nu^2 \,+\, \nu_c^2 \;=\; \gamma_p(1-\gamma_p)\sigma_x^2 \,+\, \nu_c^2.$$

Thus obtaining accurate numerical predictions for the distribution of bids is much simpler in this case than in the case of the baseline model.

Among the implications that follow in this special case: for all values of p, one obtains the prediction that

$$\mathrm{E}[\log(WTP/EV)\,|p, X] \;=\; -\frac{3}{2}\nu_c^2$$

when X is equal to its prior mean. This means that when X is equal to its prior mean, varying p cannot change the mean $\log(WTP)$; and since (3.5) implies that $\log(WTP)$ is necessarily a Gaussian distribution with variance ν_c^2, it follows that the entire distribution of values for $\log(WTP/EV)$ must be independent of p. Thus the sign of the relative risk premium cannot change as we vary p (as asserted by the fourfold pattern of risk attitudes of Kahneman and Tversky), and indeed not even its magnitude can change. It is for this reason that this case is quite inconsistent with our data.

No response noise. This model assumes that $\nu_c = 0$, so that the DM can directly (and optimally) choose their bid C as a function of the noisy internal representation $\mathbf{r}$. In this case, the optimal bidding rule is simply $C \;=\; \mathrm{E}[pX\,|\mathbf{r}]$, as assumed in the discussion in the introduction. The formulas stated above continue to hold, but with the simplification that terms involving ν_c can be omitted.

Noisy encoding and retrieval of EV. In this model, we assume (by analogy with (refrxdist)) that the internal representation of $|EV|$ is drawn from a distribution

$$r_{ev} \;\sim\; N(\log|EV|,\, \nu_{ev}^2),$$

where the variance ν_{ev}^2 is independent of the lottery's EV. The DM's bid is then assumed to have the sign of the EV (which is recognized with perfect accuracy), and a magnitude that is drawn from a distribution

$$\log|C| \;\sim\; N(f(r_{ev}),\, \nu_c^2), \tag{D.15}$$

by analogy with (3.5). The bidding function $f(r_{ev})$ is again determined so as to minimize the objective (A.1). Computing the value of this objective requires that we specify a prior regarding the true values of the lotteries with which the DM may be presented. We suppose that the DM bidding rule is optimized for a log-normal prior,

$$\log|EV| \;\sim\; N(\mu_{ev},\, \sigma_{ev}^2),$$

where (just as in the case of our other cognitive noise models) the parameters (μ_{ev}, σ_{ev}) of the prior are the ones that maximize the likelihood of the values of EV actually used in the experiment.

For the same reason as in our derivation for the baseline model, the optimal bidding function will be of the form

$$f(r_{ev}) \;=\; \log \mathrm{E}[|EV|\,|r_{ev}] \;-\; \frac{3}{2}\nu_c^2. \tag{D.16}$$

And as above, we can use the algebra of log-normal distributions to show that under these assumptions, the posterior mean will be given by

$$\mathrm{E}\Big[|EV|\,|r_{ev}\Big] \;=\; \exp\Big((1-\gamma_{ev})\bar{\mu}_{ev} \;+\; \gamma_{ev}\cdot r_{ev}\Big), \tag{D.17}$$

where we define

$$\gamma_{ev} \;\equiv\; \frac{\sigma_{ev}^2}{\sigma_{ev}^2 + \nu_{ev}^2} \;<\; 1, \qquad \bar{\mu}_{ev} \;\equiv\; \mu_{ev} \;+\; \frac{1}{2}\sigma_{ev}^2.$$

Equations (D.15)–(D.17) then completely specify the predicted log-normal distribution of bids implied by any internal representation r_{ev}.

From this, we obtain the prediction that $m(p, X)$ should be a log-linear function of the form (1.2), with

$$\alpha_p \;=\; (1-\gamma_{ev})[\bar{\mu}_{ev} - \log p] \;-\; \frac{3}{2}\nu_c^2, \qquad \beta_p \;=\; -(1-\gamma_{ev}),$$

and that

$$v(p, X) \;=\; \gamma_{ev}^2\nu_{ev}^2 \;+\; \nu_c^2$$

for all $(X;\,p)$. Thus this model, like the baseline model, predicts that $m(p, X)$ should be an affine function of $\log|X|$, with a slope between 0 and -1, that is independent of the sign of X. It also predicts that α_p should be monotonically decreasing as a function of p, rising sharply for the smallest values of p, in qualitative accordance with our estimated coefficients for the atheoretical bounded symmetric affine model. However, like the model with exogenous noise (or the model with noiseless retrieval of monetary payoffs), it predicts that β_p should be the same for all p, rather than becoming more negative for small values of p, as in our data. And it predicts a sharper rate of increase in α_p for small values of p than does the baseline model: this model predicts that α_p should equal a constant minus $\log p$, while the baseline predicts that it should equal a constant plus $\mathrm{E}[\log \mathrm{E}[p|r_p]\,|p]$ minus $\log p$. Finally, this model predicts that the trial-to-trial variability of bids (in percentage terms) should be independent of both p and $|X|$; this means that (unlike the baseline model) it fails to capture the way in which bids are more variable for low values of p.

The best-fitting (maximum-likelihood) parameter estimates for each of the alternative models, when fitted to the moments of the bidding behavior of the average subject reported in Figures 3-4, are given in Table 4.[80] (The parameter value reported for the cost of precision is the value of $\tilde{A}$, rather than A, since it is the former parameter that can be directly recovered from subjects' bidding behavior.) The degree to which the model fits in each case is indicated

[80] The method of parameter estimation is discussed below in Appendix section E.

Variants of the Baseline Model			
model	$\tilde{A}$	ν_z	ν_c
baseline model	0.002	1.60	0.24
no payoff noise	0	1.61	0.25
no probability noise	0.019	0	0.50
no response noise	0.004	1.75	0
Model with Exogenous Precision			
	ν_x	ν_z	ν_c
	0.16	1.61	0.24
Model with Noisy Retrieval of EV			
	ν_{ev}		ν_c
	0.95		0.24

Table 4: Parameter estimates for the alternative cognitive noise models, for which model comparison statistics are given in Table 2 of the main text.

by the log-likelihoods and BIC statistics reported in Table 2 of the main text. The table also repeats the parameter estimates for the baseline model (with all three types of cognitive noise, and endogenous precision of encoding of monetary payoffs), for purposes of comparison with the other models.

E Maximum Likelihood Parameter Estimation

E.1 Likelihood of the Individual-Trial Data

Let y_i be the observed value on any trial i of the variable $\log(WTP/EV)$. The log-likelihood of the data $\{p_i, X_i, y_i\}$ can be expressed in the form

$$\text{LL} \;=\; \sum_i \left[L_1(p_i, X_i) \;+\; L_2(y_i\,|p_i, X_i)\right], \tag{E.1}$$

where the sum is over the trials in the data set, indexed by i. For each trial, the contribution $L_1(p_i, X_i)$ is the log of the likelihood of the subject's being presented with lottery (p_i, X_i) according to the prior; and $L_2(y_i\,|p_i, X_i)$ is the log of the conditional likelihood of the (scaled) response y_i, given lottery (p_i, X_i), under a given parametric model of bidding behavior. In our atheoretical models, the parts L_1 and L_2 are each functions of different sets of parameters: the parameters of the priors matter only for L_1, while the behavioral parameters matter only for L_2. But in our optimal bidding model, instead, the conditional likelihoods L_2 also involve the parameters of the prior, in the way explained in Appendix section D. (In the case of the comparisons between alternative atheoretical models in Table 1, the addition of the L_1 term to our definition of LL has no effect on our model comparisons, since it simply adds the same constant to the value of LL on each line.)

We can write (E.1) in the form

$$\text{LL} \;=\; \sum_j N_j L_j, \tag{E.2}$$

where the sum is over the different lotteries (indexed by j) used in the experiment, N_j is the number of trials involving lottery j, and L_j is the average contribution to the log likelihood from the trials involving that lottery. Each term L_j depends only on the data for trials $i \in I_j$, the set of trials on which $(p_i, X_i) = (p_j, X_j)$. Thus L_j depends only on p_j, X_j, and the bids $\{WTP_i\}$ for trials $i \in I_j$. We can also further decompose each of the terms LL_j in the same way as in (E.1), writing

$$L_j \;=\; L_1(p_j, X_j) \;+\; L_{2,j}, \tag{E.3}$$

where

$$L_{2,j} \;=\; \frac{1}{N_j} \sum_{i \in I_j} L_2(y_i \,|p_j, X_j).$$

The L_1 terms are the same for all of the models that we consider in this paper. Our specifications (2.2) and (3.3) for the prior imply that

$$L_1(p_j, X_j) \;=\; -\frac{1}{2}\left(\frac{\log|X_j| - \mu_x}{\sigma_x}\right)^2 \;-\; \log(\sqrt{2\pi}\sigma_x) \;-\; \log(2\sqrt{3}\sigma_z), \tag{E.4}$$

for any p_j such that

$$\mu_z - \sqrt{3}\sigma_z \;\leq \log\frac{p_j}{1-p_j} \;\leq\; \mu_z + \sqrt{3}\sigma_z. \tag{E.5}$$

(Here we have omitted certain additive terms in (E.4) that are independent of the assumed parameter values; these terms have no effect on our judgments about the relative value of LL under different parameter values, and hence no effect on our maximum-likelihood parameter estimates or our model-comparison statistics.)

If p_j instead falls outside the interval (E.5), i.e., outside the support of the prior (3.3), given the assumed parameter values, then the prior probability of such an observation is zero, and $L_1(p_j, X_j) = -\infty$. Hence in our search for maximum-likelihood parameter values, we can impose as a constraint that the parameters of the prior must satisfy

$$\mu_z - \sqrt{3}\sigma_z \;\leq \min_j \log\frac{p_j}{1-p_j}, \qquad \mu_z + \sqrt{3}\sigma_z \;\geq \max_j \log\frac{p_j}{1-p_j},$$

where the minimum and maximum are over the set of probabilities used in the experiment.[81] Subject to these constraints, we find values of the parameters that maximize the function LL, using expression (E.4) for the L_1 terms.

E.2 The Likelihood Expressed in Terms of First and Second Moments of the Distributions of Bids

In each of the atheoretical characterizations of the data considered in Table 1, we assume a distribution of bids for the lottery (p_j, X_j) of the form

$$y_i \;\sim\; N(m_j,\, v_j) \tag{E.6}$$

[81] As shown in Table 3, these minimum and maximum probabilities are 0.05 and 0.95 respectively.

on each trial $i \in I_j$; the models differ only in the restrictions that they place on the possible values of the parameters $\{m_j,\, v_j\}$. In the case of any model of this kind, the average contribution of each trial involving lottery j to the conditional log-likelihood of the data is then given by

$$L_{2j} \;=\; -\frac{1}{2v_j}\left[\hat{v}_j \,+\, (\hat{m}_j - m_j)^2\right] \;-\; \frac{1}{2}\log(2\pi v_j), \tag{E.7}$$

where we define the sample mean and variance of the data as

$$\hat{m}_j \;\equiv\; \frac{1}{N_j}\sum_{i\in I_j} y_i, \qquad \hat{v}_j \;\equiv\; \frac{1}{N_j}\sum_{i\in I_j}(y_i - \hat{m}_j)^2.$$

Note that in (E.7), the quantities m_j, v_j are parameters of the model (the values of which are estimated to fit the data), while the quantities $\hat{m}_j, \hat{v}_j$ are data moments. Given the data, the MLE estimates for the parameters (in the absence of any further restrictions) will depend only on these moments of the data, and are equal to[82]

$$m_j \;=\; \hat{m}_j, \qquad v_j \;=\; \hat{v}_j.$$

Thus in the case of any model parameters $\{m_j,\, v_j\}$, the value of the log-likelihood LL can be computed from the data moments $\{\hat{m}_j,\, \hat{v}_j\}$, using equations (E.2) – (E.4) and (E.7).

In the results reported in Table 1, the parameters of each of the various atheoretical statistical models are fit to the data moments of a fictitious "average subject." For each lottery j, we define $\hat{m}_j^{avg}$ as the median value of $\hat{m}_j$ across the various subjects who bid on lottery j, and $\hat{v}_j^{avg}$ as the median value of $\hat{v}_j$ across these same subjects. (These are the data moments plotted in Figures 3 and 4.) In order to compute the log likelihood for any model parameters $\{m_j,\, v_j\}$ using equations (E.2) – (E.4) and (E.7), using $\{\hat{m}_j^{avg},\, \hat{v}_j^{avg}\}$ for the data moments in (E.7). For the quantity N_j in (E.2), we use N_j^{avg}, the effective number of observations of bids on lottery j by the average subject. This is defined as

$$N_j^{avg} \;\equiv\; \frac{1}{H_j}\sum_h N_j^h,$$

where H_j is the number of subjects bidding on lottery j.[83]

The MLE estimates of the parameters of the cognitive noise models are chosen in a similar way, to maximize the log likelihood of the average-subject data. The exact solution to the optimal bidding model does not imply that a DM's bids on a given lottery should be drawn from a log-normal distribution, as specified in (E.6); while (D.3) implies a log-normal distribution of bids conditional on the internal representation $\mathbf{r}$, when we condition on the true lottery characteristics (as in our computation of the data moments) rather than on the unobserved internal representation, the predicted distribution should instead be a mixture of log-normal distributions. For purposes of model fitting, however, we use a Gaussian approximation to the model predictions, according to which y_i should have a

[82]This explains our notation for the data moments: $\hat{m}_j$ is the MLE estimate of the parameter m_j, and $\hat{v}_j$ is the MLE estimate of the parameter v_j.

[83]Note that this is not the same for all lotteries j. The value of H_j varies between 5 (in the case of lotteries with $p = 0.3$ or 0.7) and 22 (in the case of lotteries with $p = 0.1$ or 0.9); see Table 3 above.

log-normal distribution as specified in (E.6), the parameters of which are given by the mean and variance of $\log y_i$ predicted by the optimizing model. Using this approximation, we can compute an approximate likelihood of the data under any assumed model parameters, simply on the basis of data for the first and second moments $\{\hat{m}_j^{avg}, \hat{v}_j^{avg}\}$.[84]

The MLE estimates of the parameters of our various theoretical models are also obtained by maximizing an approximate likelihood function calculated in this way. The reported values of LL and BIC are then based on the maximized value of the approximate likelihood function. Finally, the value of LL/N reported in Table 5 for the cognitive noise model with a single average subject actually divides LL by $N^{avg} = N/H$, the average number of bids per subject, where N is the total number of trials on which non-zero bids are submitted and H is the number of subjects in the group from which we select the median moments.

F Heterogeneity of Subject Responses

In the main text, we characterize the responses of an "average subject," and compare these to the predictions of a variety of cognitive noise models. Here we provide additional information about the heterogeneity of individual subjects' responses.

F.1 Heterogeneous Stake-Size Effects

In section 1 of the main text, we have described only the behavior of an average subject, by presenting for each lottery the median values of the individual subjects' mean $\log WTP$ and s.d.$[\log WTP]$. It is worth noting, however, that the bidding of the many of the individual subjects is at least qualitatively similar to the patterns shown in Figures 3 and 4.

We can reduce the number of statistics required to summarize the behavior of each of our subjects if, for each value of p faced by that subject, we report the coefficients (α_p, β_p) of a linear regression of the form (1.2). That is, we fit a symmetric affine model to the data for each of our 28 subjects, but allow the coefficients $\{\alpha_p, \beta_p\}$ and the residual variance v_j for each lottery to differ for each subject. The estimated regression coefficients for the different subjects are then plotted as functions of p in Figure 11. (Dashed lines connect the points corresponding to the coefficients for a given subject but for different values of p.)

While there is clearly variation in lottery valuations across subjects, we note that the general patterns of behavior identified in the data for the average subject hold also at the individual level, in most cases. In particular, we find stake-size effects ($\beta_p \neq 0$) in the case of the majority of our subjects, and in most cases we find that $-1 \leq \beta_p \leq 0$ holds (or is not clearly rejected) for all p. This is especially true in the case of the subjects who undertook 640 trials over the session; in this group β_p remains well below zero for the majority of subjects over the entire range of values for p.

We also observe a fairly consistent pattern across subjects in how both coefficients vary with p: α_p is larger (meaning a greater tendency toward risk-seeking in the gain domain and risk-aversion in the loss domain) for smaller values of p, and β_p is more negative (meaning

[84] Use of this approximation is desirable, not simply as a way of simplifying our numerical solution for the likelihood, but because we have only defined the first and second moments of the "average subject data" — we don't have a complete sample of bids by the fictitious "average subject."

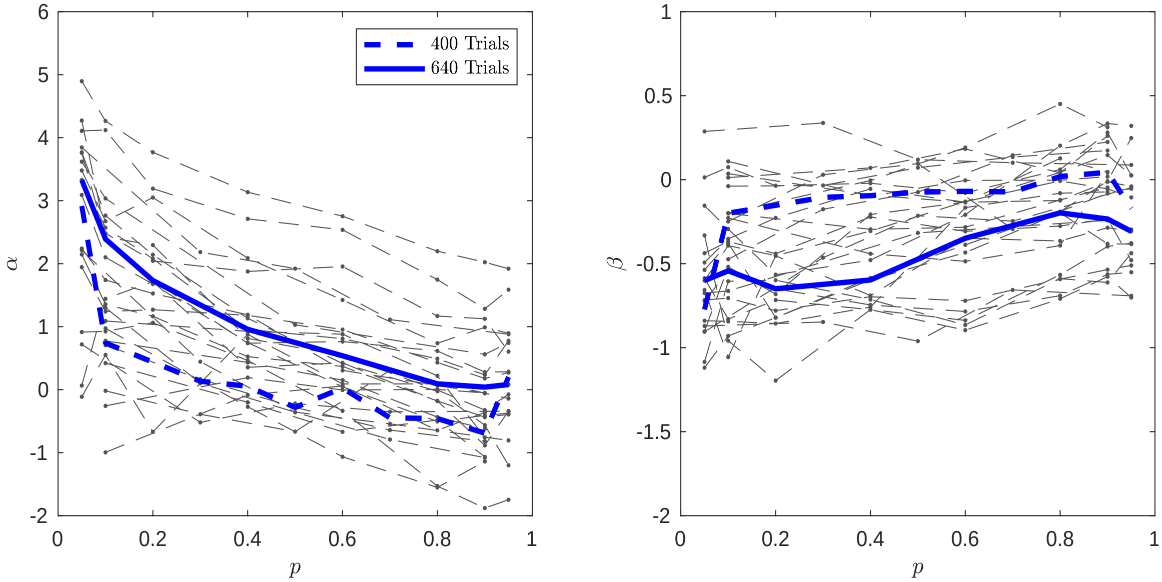


Figure 11: The coefficients $\{\alpha_p, \beta_p\}$ of the best-fitting symmetric affine model, estimated separately for each of our 28 subjects, and plotted as a function of p for each subject. The heavy curves indicate the median coefficients for each of two groups of subjects: the 13 who each completed 400 trials, and the 15 who each completed 640 trials.

more pronounced stake-size effects) for smaller values of p. In the main text, we explain why these are implications of our baseline cognitive noise model.

Finally, we note a consistent pattern in the difference between the responses of subjects who evaluated different numbers of lotteries.[85] For all values of p, α_p tends to be larger, and β_p more negative, in the case of the subjects who undertook 640 trials relative to those who undertook only 400 trials. This suggests a possible effect of time pressure or fatigue, not simply on the variability of responses, but on a subject's average valuations. Such an effect is puzzling, if we think that subjects are reporting (though perhaps with error) valuations about which they are clear, given the specified features of the lottery; it instead has a natural explanation if we suppose that subjects' decision rules adapt in a value-maximizing way to the presence of cognitive noise.

F.2 Dependence of Model Parameters on the Number of Trials

In the main text, we have fit the parameters of our cognitive noise models to the data moments for an "average subject," but we have seen from Figure 11 that there is heterogeneity in subjects' bidding behavior. Such heterogeneity is not necessarily inconsistent with the hypothesis of an optimal bidding rule, however, if we suppose that the cognitive noise parameters need not be identical for all subjects. As an illustration of this, we estimate the parameters of our baseline cognitive noise model separately for two different "average

[85] As explained in the Appendix, section A.3, the two groups do not differ only in the number of questions that they were required to answer (which might have resulted in differences in the degree of fatigue or concentration). The groups also differ in the values of p used in the lotteries that they evaluated, though both groups faced both small and large values of p.

Parameter Estimates: Baseline Cognitive Noise Model					
data	$\tilde{A}$	ν_z	ν_c	LL	LL/N
400-trial avg. subject	0.000004	1.19	0.29	-1202.6	-3.015
640-trial avg. subject	0.0078	2.28	0.28	-1995.8	-3.171
both average subjects	0.0021	1.75	0.27	-3249.8	-3.161
single average subject	0.0017	1.60	0.24	-1602.5	-3.068
Alternative Models of Both Average Subjects					
model	LL			BIC	K
common parameters	-3249.8			6534.2	1
separate parameters	-3198.4			6459.0	2.1×10^{16}

Table 5: Alternative estimates of the cognitive noise parameters for the baseline model, depending which average subjects' bidding behavior the model is required to explain. The upper part of the table presents the parameter estimates and a measure of the model's ability to fit each set of behavioral moments. The bottom part of the table compares two alternative uses of the model to explain the joint behavior of the 400-trial and 640-trial average subjects: one in which separate parameters are estimated for each average subject, and another in which the parameters are constrained to be the same for both.

subjects," one based on the 13 subjects who each evaluated 400 lotteries, and the other based on the 15 subjects who each evaluated 640 lotteries.

The upper part of Table 5 shows how the estimated cognitive noise parameters differ across four possible versions of our model: a model fit only to the data of the 400-trial "average subject"; a model fit only to the data of the 640-trial "average subject"; a model fit to the data moments of the two "average subjects" together, but with a single set of parameters required to explain the behavior of both; and a model fit to the data moments of a single overall "average subject" (the baseline model in Table 2). For each of these estimation exercises, the maximized LL of the data moments is reported. The final entry in each row reports the value of LL divided by N, the number of observations in that dataset. This allows us a measure of the degree to which the optimizing model is able to fit the average subjects' behavior that is comparable across the different cases, despite the differing number of observations that are used to compute LL in the different cases.

The same method as explained in section F is used in Table 5 to compute MLE parameter estimates (and values for LL and BIC) based on the data for these alternative "average subjects." For example, in the case of the 640-trial average subject, the lotteries j for which the moments are computed are only the 80 lotteries used for subjects in group 5 (the 640-trial subjects), and the sums are only over the subjects h that belong to group 5.[86] In (E.2), N_j is now understood to mean $\sum_h N_j^h$, where the sum is only over the subjects in group 5. Finally, in calculating N_j^{avg}, we use the number of subjects in the 640-trial group for the value of H_j;[87] and in computing LL/N, we use a value N^{avg} that divides the total number

[86] For the different groups of subjects, and the lotteries evaluated by each group, see Appendix section A.3 above.

[87] Thus $H_j = 12$ in the case of the 640-trial average subject, for each of the lotteries on which group 5 bid.

of trials by the 640-trial subjects by the number of such subjects.[88]

In the case of the 400-trial average subject, we similarly compute moments only for the 100 lotteries that are evaluated by at least some of the subjects in groups 1-4 (the 400-trial subjects), and for each lottery j of this kind, we sum only over the subjects h in the groups that evaluate lottery j. For each lottery j in this set of 100 lotteries, H_j is the number of 400-trial subjects who evaluate lottery j (which varies across lotteries). And in computing LL/N, we use a value N^{avg} that divides the total number of non-zero bids by the 400-trial subjects by the number of such subjects.[89]

We observe that the parameter values that best fit the behavior of the 640-trial average subject are fairly different from those that best fit the behavior of the 400-trial average subject. As mentioned in the main text, the 640-trial average subject has a much larger cost of precision in magnitude representation (and hence less precise representations of the monetary payoffs), and noisier internal representations of the probabilities as well, though the degree of response noise is similar for both. Moreover, the best-fitting parameters for either of the two groups are fairly different from those estimated when we require a single set of parameters to fit both average subjects (third line of the table), or when we fit the model to an average subject that pools the data from both groups of subjects (the bottom line). The cognitive noise model fits better (in the sense of achieving a high value of LL/N) to moments of the 400-trial average subject than to the data moments of the average subject when all 28 subjects were considered as a single group.

The bottom part of Table 5 demonstrates the value of allowing for heterogeneity in the parameters of the two groups through a formal model comparison. We consider two possible quantitative models of the 260 data moments consisting of the 100 data moments of the 400-trial average subject and the 160 moments of the 640-trial average subject. In one model (the "separate parameters" model), we fit the model separately to the moments of each of the two average subjects; the best-fitting parameter values for each subject are the ones shown on the first two lines of the upper part of the table. The LL for this model is just the sum of the LLs shown on those two lines. In the other model (the "common parameters" model), we instead require the values of the parameters to be the same for both average subjects; the best-fitting parameter values for this exercise are shown on the third line of the upper part of the table. The LL for this model is also taken from the third line in the upper part of the table. Since the two models involve different numbers of free parameters, we compare their degree of fit using the BIC rather than the LL alone. Because the "common parameters" model is more parsimonious, the difference in the BICs of the two models is not as great as twice the difference in their LLs. Nonetheless, the "separate parameters" model fits the data better, even using the BICs as the basis for our judgment. The implied Bayes factor in favor of the "separate parameters" model is greater than 10^{16}.

[88] Note that the number of bids N^{avg} by the 640-trial average subject is not 640, because of the trials on which subjects in group 5 decline to bid. For the 640-trial average subject, N^{avg} is actually only equal to 629.33. This is why in Table 5, the number given for LL/N is not equal to the number given for LL divided by 640.

[89] Because the fraction of zero bids is smaller in the case of the 400-trial subjects, as discussed below, this results in $N^{avg} = 398.85$, a number only slightly less than 400.

Stochastic Prospect Theory				
value function	prob. weighting	LL	BIC	LL(o.o.s.)
power law	linear	-1878.3	3775.4	-2189.4
power law	TK92	-1653.9	3332.9	-1965.1
power law	Prelec	-1627.1	3285.4	-1941.2
logarithmic	TK92	-1654.3	3333.6	-1965.5
logarithmic	Prelec	-1626.1	3283.5	-1940.4
Cognitive Noise Model				
baseline model		-1602.5	3236.3	-1917.6

Table 6: Model comparison statistics for the fit of several stochastic versions of prospect theory to the distributions of bids of our average subject. The final column gives the log-likelihood of the data under an out-of-sample prediction exercise. The bottom line presents the corresponding statistics for our baseline model, for comparison.

G Comparison with the Fit of Prospect Theory

As a benchmark for judging the degree of fit of our cognitive noise model, it is useful to compare the fit to our data of another kind of parametric model (albeit without a foundation in optimization), namely prospect theory (PT). As is well known, PT provides an explanation for the fourfold pattern of risk attitudes documented by Kahneman and Tversky, and it can be specified so as to allow for stake-size effects as well. Like our baseline model, some quantitative versions of PT involve as few as three free parameters: one to specify the degree of nonlinearity of the "value function" applied to gains or losses, one to specify the degree of nonlinearity of the "weighting function" that modifies the probabilities of the different outcomes, and one to specify the degree of random error in subjects' individual responses (Stott, 2006). We begin by reviewing the alternative quantitative specifications of PT that we consider, and then discuss the model-comparison statistics implied by each of them.

Table 6 reports the log likelihood of the average-subject data, and the corresponding BIC statistic, for several possible stochastic versions of PT, using parametric specifications of the value function and weighting function that have been popular in the empirical literature. In each case, we make PT stochastic (allowing us to calculate a likelihood for our experimental data) by assuming a multiplicative response error (3.5), just as in our baseline model; but now the bidding intention $f(\mathbf{r})$ is replaced by the valuation of the lottery $(X; p)$ implied by a deterministic version of PT. In all of the versions of PT considered in Table 6, we assume (for the sake of parsimony) that the same value function and weighting function apply in both the gain and loss domains; asymmetric versions of PT are considered further below. We begin by explaining the precise specification of the models considered in Table 6.

G.1 Stochastic Versions of Prospect Theory

In all of the versions of PT that we discuss, we assume that bids are drawn from the distribution (3.5), except that now f is a function of the objective data (p, X) rather than of a noisy internal representation. Here $f(p, X)$ is the logarithm of the absolute value of the bid $\bar{C}$ implied by deterministic PT; thus we modify PT by multiplying the deterministic prediction

$\bar{C}(p, X)$ by a log-normally distributed response error.[90] The deterministic prediction is the monetary amount $\bar{C}$ such that

$$V(\bar{C}; 1) \;=\; V(X; p), \tag{G.1}$$

where PT assigns a value (in non-monetary units) of

$$V(X; p) \;\equiv\; w(p) \cdot v(X)$$

to a random prospect offering the monetary amount X with probability p (and zero otherwise). Thus $\bar{C}$ is the amount of money such that, according to (deterministic) prospect theory, a DM should be indifferent between receiving $\bar{C}$ with certainty and receiving X with probability p.

We consider a variety of different specifications for the value function $v(X)$ and the probability weighting function $w(p)$, each of which has been popular in the empirical literature. In the *symmetric* versions of PT that we consider, we impose the restriction that $v(-X) = -v(X)$, in which case PT (like our noisy coding model) predicts that (except for the sign of the bids) the distribution of bids for any values of p and $|X|$ are the same in the case of both lotteries involving gains and lotteries involving losses. More generally, one might allow the function to be asymmetric. In the empirical fits reported in Tversky and Kahneman (1992), an asymmetric *power law* function is assumed: for values of X with either sign, it is assumed that

$$v(X) \;=\; \text{sign}(X) \cdot |X|^{\alpha} \tag{G.2}$$

for some $0 < \alpha \leq 1$, with the value of α allowed to differ depending on the sign of X.[91] The cases called "power law" in Table 6 assume a *symmetric* power law function: a function of the form (G.2), with the same value of α regardless of the sign of X.[92]

The "power law" specification (G.2) has been very popular in empirical implementations of PT, as discussed by Stott (2006). But in the case of a power-law value function, equation (G.1) reduces to

$$\bar{C} \;=\; \left(\frac{w(p)}{w(1)}\right)^{\frac{1}{\alpha}} X,$$

which implies that the median value of WTP/EV (i.e., $\bar{C}/pX$) should be a function of p, independent of the value of X. Thus there should be no stake-size effects under this version of prospect theory. Other choices of value function can instead allow for stake-size effects, as discussed by Scholten and Read (2014). The most popular choice in the empirical literature

[90] As reviewed in Stott (2006), empirical implementations of PT often make the theory stochastic by *adding* a random term to $\bar{C}$ rather than assuming a multiplicative valuation error. However, the studies reviewed there generally model choices between pairs of lotteries, rather than elicited certainty equivalent values, as here. The use of a multiplicative response error specification makes the stochastic PT models that we consider more comparable to the variants of our cognitive noise model.

[91] Tversky and Kahneman also allow for a positive multiplicative factor different from 1, that can differ depending on the sign of X. (They argue for a larger multiplicative factor in the case of losses, reflecting loss aversion.) However, such a multiplicative factor has no consequences for the predicted valuations of prospects that involve payoffs that are all of the same sign, as in the case of the lotteries used in our experiment. Hence we can without loss of generality assume a multiplicative factor of 1, in the case of either gains or losses. We can instead separately identify the value of α for the cases of gains and losses respectively.

[92] We have also estimated variant models in which α is allowed to differ for gains and losses (results not reported here), but do not find an improvement in fit (increase of LL) sufficient to justify the additional free parameter, if the BIC is used to judge model fit. Hence we only report results for symmetric models here.

value function	prob. weighting	#params	α	γ	δ	ν_c
power law	linear	1	1			0.55
power law	TK92	2	1	0.540		0.23
power law	Prelec	4	0.849	0.519	0.838	0.21
logarithmic	TK92	3	0.001	0.542		0.23
logarithmic	Prelec	4	0.043	0.520	0.839	0.21

Table 7: Maximum-likelihood parameter estimates for the five stochastic versions of prospect theory referred to in Table 6. The column "#params" indicates the number of free parameters penalized when computing the BIC statistics reported in Table 6.

focusing on stake-size effects has been the logarithmic specification advocated for example by Bouchouicha and Vieider (2017),

$$v(X) \;=\; \mathrm{sign}(X) \cdot \log(1 + \alpha X) \tag{G.3}$$

for some $\alpha > 0$. The cases called "logarithmic" in Table 6 assume a function of this kind with the same value of α for both gains and losses. We consider this specification in order to give PT as good a chance as possible to fit the stake-size effects that we find.

For the weighting function $w(p)$, the simplest case that we consider (called "linear" in Table 6) assumes that $w(p) = p$; in this case, PT is equivalent to a version of expected utility maximization, in which however the nonlinear utility function is applied to the *change* in wealth from an individual gamble rather than to the DM's overall wealth. (This version of expected utility theory does not correspond to the original proposal of Bernoulli, 1954 [1738], but is one that is commonly used in experimental studies of decision making under risk.) Tversky and Kahneman (1992) instead consider a one-parameter family of nonlinear weighting functions,

$$w(p) \;=\; \frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^{1/\gamma}}, \tag{G.4}$$

for some $0 < \gamma \leq 1$. (Note that in the limiting case $\gamma = 1$, this reduces to the linear specification.[93] For values $0 < \gamma < 1$, the value function has an "inverse S shape" of the kind hypothesized by Kahneman and Tversky with limiting values $w(0) = 0$ and $w(1) = 1$.) As in the case of the value function, Tversky and Kahneman allow the parameter γ to be different for gains and losses; the "TK92" weighting function referred to in Table 6 is instead the symmetric case in which (G.4) holds with the same value of γ regardless of the sign of X.

A variety of other nonlinear probability weighting functions have been proposed in the literature, as reviewed by Stott (2006). Among these, the simple family (family of functions with two free parameters or fewer) that fits our data best is the two-parameter family proposed by Prelec (1998), in which

$$w(p) \;=\; \exp(-\delta(-\log p)^\gamma), \tag{G.5}$$

[93] We nonetheless consider the linear case separately in Table 6, because assuming linearity *a priori* reduces the number of free parameters of the model. It would thus be possible for the model that imposes $w(p) = p$ to fit better, according to the BIC criterion, than the best-fitting member of the family of models that assume (G.4).

for parameter values $0 < \gamma, \delta \leq 1$. (This family also nests the linear specification in the case that $\gamma = \delta = 1$, while it implies an "inverse S shape" if γ and δ are both less than 1. Prelec derives this family of functions from an attractive set of axioms.) We present results for this alternative (called "Prelec" in Table 6) in order to show the case (among those that we have investigated) in which a stochastic version of PT is most successful in fitting our data.

G.2 Maximum-Likelihood Estimates and In-Sample Model Comparisons

The parameter values that best fit the data from our average subject are shown in Table 7, for each of several versions of PT. These parameter values are then used to compute the log-likelihood of the data, and the implied BIC statistic, reported in the columns labeled "LL" and "BIC" in Table 6.

For each combination of functional forms for the value function and probability weighting function, Table 7 shows the best-fitting parameter values when the same functions are used for both lotteries involving gains and those involving losses. In the case that the power law value function (G.2) is combined with either a linear probability weighting function (the expected utility case) or the TK92 weighting function, we find that the best-fitting parameter value, subject to the constraint that $\alpha \leq 1$, is $\alpha = 1$ (a linear value function). This means that the model on the first line of Table 7 is one in which both the value function and weighting function are linear; this corresponds to EV maximization, but with a (multiplicative) random response error. The single free parameter to estimate in this case is the value of ν_c. In the model on the second line, the constraint also binds, so that the predicted mean value of $\log(WTP/EV)$ for each value of p is due solely to the nonlinearity of the probability weighting function; the best-fitting value of γ is the one that best fits the implied function $w(p)$ to a graph of WTP/X as a function of p, of the kind shown in Figure 1. This model has two free parameters to estimate: the weighting-function curvature parameter γ and the response noise parameter ν_c. When we instead allow the more flexible Prelec two-parameter family of weighting functions (G.5), the best-fitting value of α is somewhat less than 1, though the curvature of the best-fitting value function is still not severe.

If we instead assume the logarithmic family (G.3) of value functions, our conclusions about the best-fitting probability weighting functions are not much affected. (Compare the estimated values for γ on lines 2 and 4, or the estimated values for the Prelec parameters (γ, δ) on lines 3 and 5.) In fact, when we pair the logarithmic value function with the TK92 weighting function, the best-fitting value of α is near zero, meaning that the value function is estimated to be essentially linear (just as on line 2). When we instead pair this value function with the Prelec weighting function, the optimal value function again has more curvature; and the value function implied by the parameter value on line 5 is not shaped quite the same way as the one implied by the parameter value on line 3. The difference matters for the predicted stake-size effects; but it does not much affect the best-fitting parameter values for the probability weighting function. The consequences for in-sample model fit are shown in Table 6 above.

Comparison of the second line of Table 6 with the first shows that allowing for nonlinear probability weighting of the kind proposed by Tversky and Kahneman (1992) improves the fit of the model enough to more than offset the penalty for the additional free parameter. The predictions of the best-fitting model using the functional forms proposed by Tversky and

Kahneman (1992) are illustrated in row (a) of Figure 9.[94] The model predicts no stake-size effects of the kind observed in our data, though the nonlinear weighting function allows the model to capture the fact that $\log(WTP/EV)$ is on average higher in the case of lower values of p. Comparison of the third line of Table 6 with the second shows that the more complex Prelec specification of the weighting function fits even better, again even allowing for the penalty for additional free parameters. Using a logarithmic value function instead of the power law does not improve the fit when combined with the TK92 weighting function. But when combined with the Prelec weighting function, the logarithmic value function does fit better, and in fact, the version of PT that combines a logarithmic value function with Prelec's weighting function fits our data best, according to the BIC criterion. (The predictions of this version of PT are illustrated in row (b) of Figure 9.) Finally, the bottom line of the table shows that the cognitive noise model fits better than any of the versions of PT considered here. Indeed, the difference in BIC statistics between even the best-fitting version of PT and the cognitive noise model implies a Bayes factor larger than 17 billion in favor of the cognitive noise model.

G.3 Out-of-Sample Model Comparisons: Cross-Validation

We can alternatively compare the fit of alternative models on the basis of a measure of out-of-sample predictive accuracy, using a cross-validation approach. We split our data into five sub-samples, each of which contains bids on lotteries with a single value of $|X|$, but includes data about lotteries with all 11 of the possible values of p, and lotteries involving both gains and losses. We then independently estimate the parameters of each of our theoretical models (cognitive noise models or stochastic variants of PT) five different times, using the method explained in Appendix section E; each time, the data for a different one of the values of $|X|$ are "held out" to be used as the test of out-of-sample predictive accuracy. Thus in the first such exercise, we estimate model parameters to fit a "calibration sample" in which the value of $|X|$ takes any of the four largest values; the model with these parameters is then used to compute the log-likelihood (LL) of the "validation sample" corresponding to the lotteries in which $|X|$ takes the smallest value. As discussed in section F, the parameters are chosen to maximize the LL of the moments of our "average subject," for lotteries of the kind included in the calibration sample; the out-of-sample LL computed for the validation sample is similarly computed on the basis of the held-out moments of the bids of the average subject.

We then repeat the same exercise using the moments of bids on all lotteries except those with the second-lowest value of $|X|$ as the "calibration sample," and the moments of bids on lotteries with the second-lowest $|X|$ as the "validation sample"; and so on. In this way, we obtain an out-of-sample LL for each of the five subsamples. We add these five quantities to obtain an out-of-sample LL for the complete set of moments of the average subject (the ones used for the in-sample LL measures shown in the column "LL" in Table 6). These out-of-sample measures of the log-likelihood are shown in the column of Table 6 labeled "LL (o.o.s.)." By construction, the out-of-sample LL for each model is lower than the in-sample

[94] In Figure 9, we show only one row for each model, as we consider here only symmetric models, in which the predictions are identical for lotteries involving gains or losses.

LL.

The relative size of the out-of-sample LL for different models can be used as an alternative basis for model comparisons. A higher value of the out-of-sample LL indicates that a model is more consistent with our data; and if two models have respective out-of-sample LLs of LL_1 and LL_2, where $LL_1 > LL_2$, then the likelihood ratio

$$LR \equiv \exp(LL_1 - LL_2) > 1$$

provides a measure of the factor by which one's posterior should favor model 1 over model 2, if the two models were assigned an equal prior probability of being correct. Thus the measure LR can be used in a similar way as the Bayes factor K reported in Table 1 (based on in-sample model fit).

Note that there is no need to correct for the different numbers of free parameters in our alternative models in the case of these out-of-sample comparisons: we can simply compare the values of LL for the different models, without making a correction of the kind involved in the computation of the BIC. The reason is that in each of our out-of-sample prediction exercises, none of the model's free parameters can be adjusted so as to be better fit any of the moments in the validation sample — that is, the set of moments for which the out-of-sample LL is reported. "Over-fitting" of the data in the calibration sample should be penalized by poorer out-of-sample predictive accuracy in the validation sample, without any need for an additional penalty.[95]

The out-of-sample comparisons lead to similar conclusions as model selection on the basis of the BIC: the best-fitting PT variant is the one that combines the logarithmic value function with the Prelec weighting function, but the baseline cognitive noise model still fits better than the best of the PT variants. Even for the best PT variant, the out-of-sample LLs imply a likelihood ratio of 8 billion in favor of the cognitive noise model.

G.4 Asymmetric Variants of Prospect Theory

The variants of PT considered in Tables 6 and 7 all assume that the same parameters specify the value function, the probability weighting function, and the amount of multiplicative response noise in the case of lotteries involving either gains or losses. However, many authors fitting empirical versions of PT to experimental data follow Tversky and Kahneman (1992) in fitting separate parameters to the data for lotteries involving gains and the data for lotteries involving losses. Here we consider the fit of these less-parsimoniously parameterized variants of PT as well.

Table 8 reports the same statistics as in Table 6, but for both symmetric and asymmetric variants of the models considered. (In each case, the "symmetric" model imposes the constraint that parameters are the same for either gains or losses, while the "asymmetric" model has separate parameters for gains and for losses. Thus in each case, the number of free parameters is twice as large in the case of the asymmetric model.) To keep the size of the

[95] This possibility is illustrated by the comparison in Table 8 below between the two variants of the "TK92" model. The asymmetric variant is a more flexibly parameterized version of the symmetric model, and as such must achieve a higher in-sample LL, though the BIC is higher (owing to the penalty for the additional free parameters). But the out-of-sample LL is *lower* in the case of the more flexibly-specified model.

Stochastic Prospect Theory				
model	symmetry?	LL	BIC	LL(o.o.s.)
TK92	symm.	-1653.9	3332.9	-1965.1
TK92	asymm.	-1653.1	3341.0	-1965.5
log-Prelec	symm.	-1626.1	3283.5	-1940.4
log-Prelec	asymm.	-1620.4	3292.2	-1937.6
Cognitive Noise Model				
baseline	symm.	-1602.5	3236.3	-1917.6
baseline	asymm.	-1598.1	3242.1	-1912.9

Table 8: Model comparison statistics for both symmetric and asymmetric versions of some of the models whose symmetric versions are compared in Table 6. The statistics reported in the three rightmost columns are the same as in Table 6.

table manageable, we report statistics for symmetric and asymmetric variants of only three of the types of models considered in Table 6: the model with the functional forms proposed in Tversky and Kahneman (1992), corresponding to the second line of Table 6, and here called "TK92"; the model that combines a logarithmic value function with the probability weighting function of Prelec (1998), corresponding to the fifth line of Table 6, and here called "log-Prelec"; and the cognitive noise model presented in section 3, corresponding to the bottom line of Table 6. Among the PT models proposed since the work of Tversky and Kahneman (1992), we present statistics here only for the "log-Prelec" model, because this is the one that fits our data best, on any of the four criteria (in-sample or out-of-sample, and imposing symmetry or not).

We see that allowing the parameters to vary between the gain and loss domains increases the log-likelihood at least slightly, for each type of model; but the increase in LL is not large enough to offset the penalty for the additional free parameters, and the BIC statistic is worse in each case for the asymmetric version of the model. (This is also true for the other models considered in Table 6, and is why we do not present statistics for the asymmetric versions of these models in the main text.) Thus from the point of view of in-sample fit, we would conclude that there is no advantage in allowing the parameters to vary between the gain and loss domains.

The conclusion is more nuanced if we instead consider out-of-sample prediction. In the case of the TK92 specification, we also find that the out-of-sample log-likelihood is lower when we estimate different parameters for gain and loss lotteries, as Tversky and Kahneman (1992) do; even though the in-sample LL is improved (necessarily), the out-of-sample LL falls, suggesting that the apparent improvement in LL in the first column reflects over-fitting to the particular dataset that is used. However, in the case of both the log-Prelec model and the baseline cognitive noise model, we find that the asymmetric versions of these models have higher log-likelihoods even out-of-sample, suggesting that there are at least small differences in the way that subjects value lotteries involving losses rather than gains.

Nonetheless, considering the asymmetric variants of these two models (or of other variants of PT) would not change the basic message of Table 6. While the log-Prelec model fits better (also out-of-sample) than the other variants of PT considered in Table 6, it does not fit as well as the cognitive-noise model, either in-sample or out-of-sample. If we use out-of-sample prediction as the basis for model comparison, we might prefer to compare the accuracy of the asymmetric versions of the two models rather than their symmetric versions. But also in this case, we would conclude that the out-of-sample LL of the baseline cognitive noise model is higher by 24.7 log points, implying a likelihood ratio greater than 50 billion in favor of the cognitive noise model. (The conclusion would thus be even stronger than when we compare the symmetric versions of the two models, as in Table 6.)

H Log-Linear Stake-Size Effects in the Data of Gonzalez and Wu (2022)

Among other studies of the valuation of simple lotteries, the study of Gonzalez and Wu (2022) is of particular interest for our purposes because, like us, they elicit certainty-equivalent values for lotteries that involve both a wide range of values of p and a wide range of monetary payoffs X. While their study also involves lotteries with more than one non-zero payoff, they consider a fairly large number of lotteries with only one non-zero payoff, like the ones used in our study; their results are directly comparable with ours on these trials. The lotteries of this kind that they use involve 8 different values of X (rather than only five, as in our study), and each of the 8 different values of X is paired with each of the 11 different values of p that they use. The probabilities that they consider also span a wider range, including values of p as small as 0.01 and as large as 0.99. The inclusion of a very small value of p is of particular interest, since we (like previous authors) find especially pronounced stake-size effects when p is small, and our theoretical model also implies that they should be especially extreme as p approaches zero.

Gonzalez and Wu (2022) also have a larger number of subjects than in our study: 47 subjects, each of whom is asked to value the same set of 165 different lotteries. There are however two disadvantages of their study relative to ours: First, they consider only lotteries involving potential gains, not lotteries involving potential losses as well. And second, they have each experimental subject value each lottery only once; thus they do not collect data on the amount of trial-to-trial variability in subjects' valuations of a given lottery.

It is nonetheless of interest to ask how the relative risk premia indicated by their data vary with p and X. We plot the median bid of their subjects for each lottery in Figure 12, using the same format as in the top row of Figures 3 and 4 of this paper.[96] As indicated by the linear regression lines included with the log-log plot in each panel, the slope of $\log(WTP/EV)$ as a function of $\log|X|$ is essentially zero for the highest values of p (all $p \geq 0.90$), but the relationship is downward-sloping for all lower values of p. Just as in our Figure 3, the most strongly negative-sloping relationships are observed for probabilities $p \leq 0.25$. The relationships are also approximately log-linear, as shown by the degree of fit of the linear regression lines.

Thus, to the extent that the data of Gonzalez and Wu (2022) can be used to address the same issues as our data, they confirm the regularities that we have noted in section 1 of the

[96] In Figure 12, there are no intervals around the median bids shown, because we have no data on trial-to-trial variation, which is what the whiskers in Figures 3 and 4 indicate.

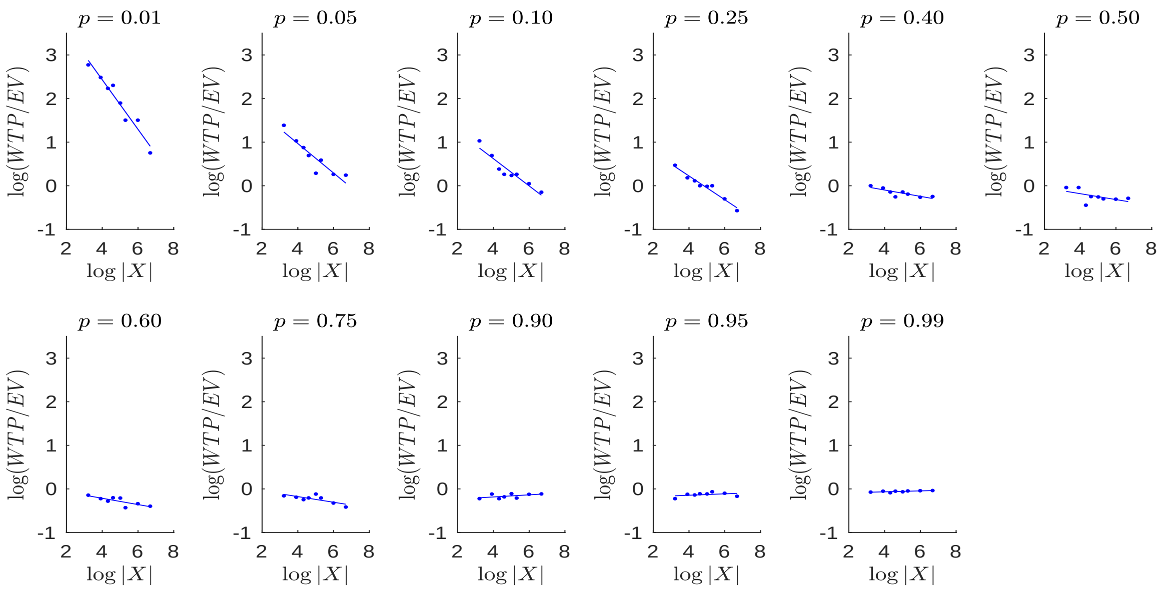


Figure 12: The value of WTP as a multiple of EV, for the median subject in the study of Gonzalez and Wu (2022). The format is the same as in the top rows of Figures 3 and 4 (here the data all refer to lotteries involving gains).

main text.[97] Indeed, they provide even stronger evidence for the effects that we document, in two important respects. First, Gonzalez and Wu consider a wider range of values for the stake size $|X|$: their largest monetary payoff (800) is 32 times as large as the smallest (25), whereas our largest payoff is only 4 times the size of our smallest; it is thus even more notable that WTP/EV appears to be a log-linear function of stake size in their data. And second, they consider a much smaller value of p (namely, 0.01) than our smallest value (0.05). They find that WTP/EV has a more strongly negative elasticity with respect to stake size when $p = 0.01$ than when $p = 0.05$ or 0.10; this is an even stronger confirmation of our conclusion that the slope becomes particularly negative for low values of p.

We can also test the fit of our baseline model to the bids of the average subject of Gonzalez and Wu. Here we treat the single value of $y_j \equiv \log(WTP_j/EV_j)$ elicited (from the median subject) for each lottery (p_j, X_j) as a single draw from the predicted distribution $N(m_j, v_j)$, where m_j and v_j depend on the values of p_j and X_j (and model parameters) in the same way as has been explained above. For any hypothesized model parameters, the log likelihood of the data is then given by $LL = \sum_j L_j$, where for each of the 88 single-nonzero-outcome lotteries j, we can again write L_j as the sum of two terms, as in (E.3). Here the expression $L_1(p_j, X_j)$ is again defined as in (E.4), while L_{2j} is now defined more simply as

$$L_{2j} \;=\; -\frac{1}{2v_j}(y_j - m_j)^2 \;-\; \frac{1}{2}\log(2\pi v_j),$$

instead of as in (E.7). We can then estimate our model parameters so as to maximize LL.

[97]This is reassuring, especially in light of the many differences in their experimental procedure: the values of X that they use are all round numbers, their subjects are not required to value as large a number of lotteries over the course of the session as ours are, certainty-equivalents are elicited using a multiple-price list in their case, etc.

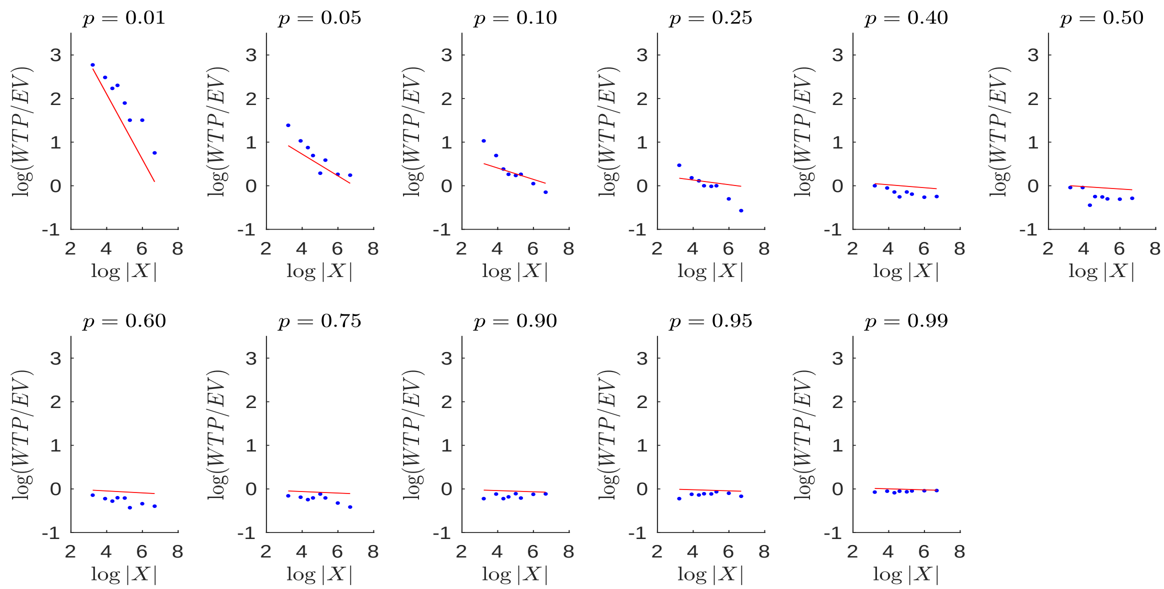


Figure 13: Fit of our baseline noisy coding model to the median bids in the study of Gonzalez and Wu (2022), treated as the bids of an average subject. The data (showing the bids) are the same as in Figure 12; the red lines show the predictions of the noisy coding model. The format is the same as in the top rows of Figures 7 and 8.

The fit of the resulting model predictions to the data plotted in Figure 12 is displayed visually in Figure 13. The dots represent the same data (the bids of the average subject) as in Figure 12, but instead of the atheoretical linear regression lines of the earlier figure, the red lines in Figure 13 plot the theoretical predictions $y_j = \alpha_p + \beta_p \log X_j$ in each panel (corresponding to a different value of p). We find that our theoretical model is broadly consistent with the data of Gonzalez and Wu (2022) as well, though the best-fitting parameter values are different than in the case of our subjects.[98]

The log likelihood of the data when the parameters of the cognitive noise model are optimized is indicated in Table 9. For purpose of comparison, the table also shows the corresponding values for the log likelihood LL and the BIC statistic for five stochastic variants of prospect theory (the same five as are compared to our model in Table 6). We see that the fit of our model to the data of Gonzalez and Wu (2022) is better than that of a number of common quantitative specifications of prospect theory, though there is at least one version of prospect theory that fits somewhat better than our baseline cognitive noise model: this is a model that combines the logarithmic value function of Bouchouicha and

[98]The maximum likelihood parameter estimates using the Gonzalez-Wu data are $\tilde{A} = 0.0004$, $\nu_z = 0.84$, $\nu_c = 0.000037$. Thus all three noise parameters are smaller when the model is fit to the bids of the average subject of Gonzalez and Wu. At least part of the difference probably reflects the fact that Gonzalez and Wu consider only lotteries involving gains; also in the case of our subjects, if we fit the model separately to the bids on lotteries involving only gains, we obtain somewhat smaller noise parameters than the ones reported in Table 4 for our baseline model: $\tilde{A} = 0.0008$, $\nu_z = 1.50$, and $\nu_c = 0.28$. Some of the difference may also reflect the fact that the subjects of Gonzalez and Wu do not have to value as large a number of lotteries, and thus may be less affected by time pressure or fatigue; recall that in Table 5 we also obtain smaller noise parameter estimates in the case of the subjects who value only 400 lotteries.

Cognitive Noise Models			
model		LL	BIC
baseline model		-320.6	663.5
Stochastic Prospect Theory			
value function	prob. weighting		
power law	linear	-408.5	830.3
power law	TK	-334.2	690.7
power law	Prelec	-323.8	674.3
logarithmic	TK	-328.5	674.8
logarithmic	Prelec	-316.6	655.5

Table 9: Model comparison statistics for the fit of several stochastic models of lottery valuation to the median bids reported by Gonzalez and Wu (2022), treated as the bids of an average subject.

Vieieder (2017) with the probability weighting function proposed by Prelec (1998), and adds a log-normal multiplicative response error to the DM's bid. It is also worth noting that the best-fitting version of prospect assumes much larger random response errors ($\nu_c = 0.058$) than does our baseline model. This is because in the case of prospect theory, any failure of the median bid to precisely fit the value implied by deterministic prospect theory must be attributed to response error; in the noisy coding model, instead, responses would be predicted to be random even in the absence of response errors (i.e., if we set ν_c equal to zero).